

SERIE A, N° 4087

N° D'ORDRE :

4938

T H E S E S

présentées à la

FACULTE DES SCIENCES
DE L'UNIVERSITE DE PARIS

pour obtenir

le grade de Docteur ès Sciences Mathématiques

par

Marcel COURTILLOT

*1ère Thèse : « Contribution à la théorie de la programmation
linéaire et de la programmation stochastique »*

2ème Thèse : Propositions données par la Faculté.

Soutenues le 24 Avril 1963 devant la Commission d'Examen

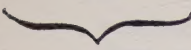
MM. FORTET Président

PISOT { Examineurs
VILLE

SERIE A, N° 4087

N° D'ORDRE :

4938



T H E S E S

présentées à la

FACULTE DES SCIENCES
DE L'UNIVERSITE DE PARIS

pour obtenir

le grade de Docteur ès Sciences Mathématiques

par

Marcel COURTILLOT

*1ère Thèse : « Contribution à la théorie de la programmation
linéaire et de la programmation stochastique »*

2ème Thèse : Propositions données par la Faculté.

Soutenues le 24 Avril 1963 devant la Commission d'Examen

MM. FORTET *Président*

PISOT { *Examineurs*
VILLE



Digitized by the Internet Archive
in 2026

https://archive.org/details/IA41550201_0024

AVANT-PROPOS

La première partie de ce travail est consacrée à la recherche de nouvelles méthodes de résolution des problèmes dits «programmes linéaires». La deuxième partie constitue une recherche dans le domaine des programmes linéaires stochastiques.

I - PREMIERE PARTIE

Nous précisons ici les réflexions qui nous ont conduit à l'élaboration de la méthode séquentielle intérieure et à la méthode du gradient itéré.

Si on considère les faces de tous ordres d'un polyèdre, on obtient un treillis (vide, sommets, arêtes, faces de dimension 2, ..., faces de dimension $n-1$, polyèdre). La méthode simplexe classique n'utilise que des cheminements du type : sommet, arête, sommet, etc... négligeant toutes les faces d'ordre supérieur. En particulier, elle néglige les cheminements possibles à l'intérieur du polyèdre. S'il est certain qu'il existe au moins un sommet du polyèdre où la borne supérieure d'une forme linéaire sur le polyèdre est atteinte, il n'est pas pour autant nécessaire de restreindre le problème à ne cheminer que de sommet en sommet.. D'ailleurs, ce mode de cheminement semble n'être utilisable que dans la classe des problèmes linéaires.

En effet, si l'on envisage de maximiser une fonction numérique (même linéaire) dans un convexe homéomorphe à un polyèdre, il n'y a pas de transposition possible.

A l'opposé des méthodes simpliciales qui n'utilisent que des cheminements superficiels du type particulier précité se trouvent les méthodes de gradient, plus naturelles et plus généralisables, qui utilisent les cheminements intérieurs. La plus connue des méthodes de gradient appliquée à la classe des programmes linéaires est celle de M. R. Frisch. Elle s'avère néanmoins incertaine quant à l'obtention d'une solution correcte.

La méthode séquentielle intérieure et la méthode du gradient itéré allient les avantages des méthodes de gradient et des méthodes simpliciales. Méthodologiquement, elles constituent un «fond continu» entre les méthodes de gradient et les méthodes simpliciales et se généralisent de façon naturelle à des problèmes de maximisation d'une fonction numérique concave dans un domaine convexe. Leur intérêt est marqué dans les domaines définis par des contraintes très nombreuses.

II - DEUXIEME PARTIE

La deuxième partie de ce travail a pour but, après avoir défini ce que l'on entend par programmes linéaires stochastiques, de donner des méthodes permettant d'une part de former des majorantes de la loi de distribution du maximum de la fonction économique et, d'autre part, de former exactement cette loi.

Par ailleurs, quelques pages ont été consacrées à ce qui pouvait être dit sur le domaine des solutions admissibles avec une probabilité donnée. L'étude se termine par la mention de deux problèmes de décision en face d'événements probabilisables.

Je veux adresser ici mes vifs remerciements à Monsieur le Professeur J. Ville qui a bien voulu diriger mes recherches, et à Monsieur le Professeur Ch. Pissot qui a bien voulu accepter de me proposer la seconde thèse.

Je suis également très reconnaissant à Monsieur le Professeur R. Fortet d'avoir bien voulu accepter de présider le Jury.

Que Monsieur C. Berline, Chef du Service de Recherche Opérationnelle de la Société des Pétroles Shell Berre, ainsi que mes collègues de ce service trouvent ici l'expression de ma gratitude pour les encouragements qu'ils m'ont prodigués lors de mes travaux de recherche.

PREMIERE PARTIE

CHAPITRE I

DEFINITIONS ET PRINCIPAUX THEOREMES

Cette première partie n'a pas la prétention d'être exhaustive mais a pour but de mentionner les principales définitions et théorèmes dont nous aurons besoin par la suite. Les théorèmes simples seront donnés sans démonstration.

1,1 - VARIETE LINEAIRE - CONES - ENSEMBLES CONVEXES

Définition 1 :

Soit X un espace vectoriel sur le corps des réels. Le sous-espace vectoriel :

$$d = [x \mid x = \lambda \alpha, \lambda \in \mathbb{R}, \alpha \in X]$$

sera dit une droite origine.

Le sous-espace vectoriel supplémentaire sera dit un *plan origine*.

Théorème 1 :

Si f est une forme linéaire définie sur X , non identiquement nulle, $P(f) = [x \mid f(x) = 0]$ est un plan origine. Inversement, tout plan origine est de cette forme.

Définition 2 :

Soit F un sous-espace vectoriel de X . Le sous-ensemble $V = F + \alpha, \alpha \in X$ est un *sous-espace affine* de X parallèle à F , ou encore *une variété linéaire* de X .

Théorème 2 :

Toute variété linéaire de X est telle que :

$$x_1, x_2 \in V, \lambda_1, \lambda_2 \in \mathbb{R}, \lambda_1 + \lambda_2 = 1 \Rightarrow \lambda_1 x_1 + \lambda_2 x_2 \in V$$

Réciproquement, toute V satisfaisant à cette relation est de la forme $F + \alpha$ donc une variété linéaire.

Définition 3 :

Une droite de X est une variété linéaire parallèle à une droite origine.

Un plan de X est une variété linéaire parallèle à un plan origine.

CONSEQUENCES : d'après le théorème 1, tout plan est de la forme $[x \mid f(x) = \lambda, \lambda \in \mathbb{R}]$ et f est une forme linéaire sur X non identiquement nulle.

Définition 4 :

K est un cône origine si $x \in K, \lambda > 0 \Rightarrow \lambda x \in K$.

Une demi-droite origine est un cône origine engendré par un seul élément.

Définition 5 :

Un ensemble C est convexe si

$$x_1, x_2 \in C, \lambda_1 + \lambda_2 = 1, \lambda_1, \lambda_2 \geq 0 \Rightarrow \lambda_1 x_1 + \lambda_2 x_2 \in C$$

CONSEQUENCES :

a) un cône convexe est tel que :

$$x_1, x_2 \in K, \lambda_1 \geq 0, \lambda_2 \geq 0, \lambda_1 + \lambda_2 = 1 \Rightarrow \lambda_1 x_1 + \lambda_2 x_2 \in K.$$

b) toute variété linéaire est convexe,

c) tout demi-espace origine est convexe

d) si C est convexe :

$$(x_1, x_2, \dots, x_n) \in C, (\lambda_1, \dots, \lambda_n) \geq 0, \sum_{i=1}^n \lambda_i = 1 \Rightarrow \lambda_1 x_1 + \dots + \lambda_n x_n \in C.$$

Théorème 3 :

Si A et B sont convexes :

$C = A \cap B$ est convexe.

Théorème 4 :

Si A et B sont convexes :

$C = A + B$ est convexe.

Théorème 5 :

Si f est une application linéaire de X dans Y , espaces vectoriels, si D est convexe $f(D) \subset Y$ est convexe.

Définition 6 :

Fermeture : soit entre 2 sous-ensembles A et $\bar{A}$ de X , une correspondance $A \rightarrow \bar{A}$ telle que :

$$A \subset \bar{A}$$

$$A \subset B \Rightarrow \bar{A} \subset \bar{B}$$

$$(\bar{\bar{A}}) = \bar{A}$$

$$\bar{\Phi} = \Phi$$

Une telle correspondance est appelée une fermeture.

CONSEQUENCES : l'intersection des convexes contenant A est une fermeture appelée fermeture convexe de A et notée $[A]$.

Théorème 6 :

Soit $A = [a_i, i \in I]$, on a :

$$[A] = \left\{ \sum_{i \in J} \lambda_i a_i \mid \forall J \text{ fini } \subset I, \sum_{i \in J} \lambda_i = 1, \lambda_i \geq 0 \forall i \in J \right\}$$

DEMONSTRATION :

$$\text{Soit } A_1 = \left\{ \sum_{i \in J} \lambda_i a_i \mid \forall J \text{ fini } \subset I, \sum \lambda_i = 1, \lambda_i \geq 0 \forall i \right\}$$

Si C est un convexe contenant A , $A_1 \subset C$ d'où $A_1 \subset [A]$

Par ailleurs $A \subset A_1$ d'où $[A] \subset [A_1] = A_1$

Par suite, $A_1 = [A]$

Théorème 7 :

Si un ensemble A est contenu dans une variété linéaire de dimension n d'un espace vectoriel X , tout $x \in A$ peut se mettre sous la forme :

$$x = \sum_1^{n+1} \lambda_i a_i, \lambda_i \geq 0, \sum_1^{n+1} \lambda_i = 1, a_1, \dots, a_{n+1} \in A$$

DEMONSTRATION :

D'après le théorème 6, tout $x \in A$ peut s'écrire sous la forme :

$$x = \lambda_1 a_1 + \dots + \lambda_p a_p, \lambda_1, \dots, \lambda_p \geq 0, \lambda_1 + \dots + \lambda_p = 1$$

Si $p \leq n + 1$ le théorème est démontré.

Supposons $p > n + 1$.

Les éléments $a_1 - a_p, a_2 - a_p, \dots, a_{p-1} - a_p, 0$ sont contenus dans un sous-espace vectoriel F de dimension $n < p - 1$. Donc il existe $\mu_1, \dots, \mu_{p-1}$, non tous nuls tels que :

$$\mu_1 (a_1 - a_p) + \dots + \mu_{p-1} (a_{p-1} - a_p) = 0$$

Si on pose : $\mu_p = -(\mu_1 + \dots + \mu_{p-1})$, on a :

$$(1) \mu_1 a_1 + \mu_2 a_2 + \dots + \mu_p a_p = 0,$$

$$\mu_1 + \dots + \mu_{p-1} + \mu_p = 0$$

Soit $\mu = \sup \left(\frac{\mu_i}{\lambda_i} \right), i \in [1, \dots, p]$, on a $\mu \geq 0$ et d'après (1)

$$x = x - 0 = \sum_{i=1}^p \left(\lambda_i - \frac{\mu_i}{\mu} \right) a_i, \sum_{i=1}^p \lambda_i - \frac{1}{\mu} \sum_{i=1}^p \mu_i = 1$$

$$\forall i : \lambda_i - \frac{\mu_i}{\mu} \geq 0 \text{ et } \exists i \in [1, \dots, p] : \lambda_i - \frac{\mu_i}{\mu} = 0$$

x s'exprime donc en fonction de $p - 1$ éléments a_i .

En opérant de cette manière $p - (n + 1)$ fois, x s'exprimera en fonction de $n + 1$ a_i au plus.

Définition 8 :

a) Soit D convexe de X , espace vectoriel et $x_0 \in D$.

Une droite d sera dite *privilégiée* par rapport à x_0 si :

$$(1) x_0 \in D \cap d$$

$$(2) \forall x \in d, \exists \lambda > 0 \Rightarrow \begin{cases} \lambda(x - x_0) \in D \\ -\lambda(x - x_0) \in D \end{cases}$$

b) Si en x_0 il n'existe aucune droite privilégiée, x_0 est dit *point extrême de D*. Autrement : x point extrême de D ne peut se mettre sous la forme $x = \lambda_1 x_1 + \lambda_2 x_2$, $x_1, x_2 \in D$, $\lambda_1 + \lambda_2 = 1$, $\lambda_1 \geq 0$, $\lambda_2 \geq 0$. L'ensemble des points extrêmes de D est noté $\tilde{D}$ et s'appelle profil de D .

c) Si en x_0 toute droite est privilégiée, x_0 est *point interne de D*.

Théorème 8 :

Les éléments $x \in \mathbb{R}^n$ appartenant aux droites privilégiées de D en x_0 forment un sous-espace vectoriel affine.

DEMONSTRATION :

Supposons $x_0 = 0$, soit Δ l'ensemble des éléments appartenant aux droites privilégiées de D en 0 , supposé non vide.

Supposons que $x_1, x_2 \in \Delta$, ceci entraîne l'existence de 2 droites privilégiées d_1 et d_2 éventuellement confondues telles qu'il existe $\lambda_1, \lambda_2 > 0$ satisfaisant à :

$$\left. \begin{matrix} \lambda_1 x_1 \\ -\lambda_1 x_1 \end{matrix} \right\} \in d_1 \cap D \quad \left. \begin{matrix} \lambda_1 x_2 \\ -\lambda_1 x_2 \end{matrix} \right\} \in d_2 \cap D$$

$$\text{Soit } \lambda = \inf(\lambda_1, \lambda_2) \text{ et } \alpha_1 = \lambda x_1, \alpha_2 = \lambda x_2$$

on a $\alpha_1, \alpha_2 \in D$. D étant convexe :

$$\frac{\alpha_1 + \alpha_2}{2} = \frac{\lambda(x_1 + x_2)}{2} \in D$$

$$\frac{(-\alpha_1) + (-\alpha_2)}{2} = \frac{-\lambda(x_1 + x_2)}{2} \in D$$

Par suite, $x_1 + x_2$ appartient à une droite privilégiée d donc $x_1 + x_2 \in \Delta$.

Par ailleurs $x \in \Delta$ entraîne de façon évidente $\lambda x \in \Delta$ pour tout $\lambda \in \mathbb{R}$.

Donc, pour $x_0 = 0$, Δ est un sous-espace vectoriel et pour $x_0 \neq 0$ Δ est un sous-espace affine.

1,2 - INTRODUCTION D'UNE METRIQUE

Si $X = \mathbb{R}^n$ et $(e_1, \dots, e_n)$ est la base canonique de $\mathbb{R}^n$, on définira la norme $\|x\| = \sqrt{\langle x, x \rangle}$

ou $\langle x, y \rangle$ est le produit scalaire $\sum_{i=1}^n x_i y_i$. La forme bilinéaire $\langle x, y \rangle$ est symétrique et

non dégénérée. Elle définit un isomorphisme de R^n sur son dual de telle sorte que :

Théorème 9 :

Si f est une forme linéaire sur R^n , il existe un élément $f \in R^n$ tel que $f(x) = \langle f, x \rangle$, de plus, $f(x)$ est continue.

CONSEQUENCES : le demi-espace $H(f, \lambda) = [x \mid f(x) \geq \lambda]$ est fermé

le demi-espace $H(f, \lambda) = [x \mid f(x) > \lambda]$ est ouvert.

Théorème 10 :

Toute variété linéaire est fermée dans R^n .

Théorème 11 : Théorème de séparation

Si D et D' sont 2 convexes disjoints D fermé et D' compact, il existe un hyperplan $P(f, \lambda) = [x \mid f(x) = \lambda]$ qui sépare strictement D et D' , c'est-à-dire :

$$\inf_{x \in D'} f(x) < \lambda < \sup_{x \in D} f(x)$$

Corollaire :

Un convexe fermé D est égal à l'intersection F des demi-espaces fermés qui le contiennent.

DEMONSTRATION :

Soit $x_0 \notin D$, x_0 est un compact, il existe un plan $P(f, \lambda)$ tel que : $f(x_0) < \lambda < \inf [f(x) \mid x \in D]$ auquel correspond le demi-espace $H(f, \lambda)$ tel que $D \subset H(f, \lambda)$, $x_0 \notin H(f, \lambda)$

x_0 ne peut donc appartenir à F , donc $F \subset D$ et comme $F \supset D$, on a $F = D$.

Définition 9 :

Un plan d'appui d'un ensemble D est un plan $P(f, \lambda)$ contenant au moins un élément de D et tel que tous les éléments de D soient d'un même côté de $P(f, \lambda)$.

Théorème 12 : Théorème du plan d'appui

Si D est un convexe compact non vide, il admet au moins un point extrême et tout plan d'appui contient au moins un point extrême de D .

DEMONSTRATION :

De façon évidente, le théorème est vrai dans R .

Supposons le vrai dans R^{n-1} , soit D un convexe compact de R^n , $P(f, \lambda)$ un plan d'appui. $P(f, \lambda) \cap D = D'$ est un fermé de D donc un compact convexe, D' contenu dans R^{n-1} admet un point extrême x_0 . Soient $x_1, x_2 \in D$ tels que : $x_0 = \lambda_1 x_1 + \lambda_2 x_2$, $\lambda_1 + \lambda_2 = 1$, $\lambda_1 \lambda_2 > 0$. On ne peut avoir $[x_1, x_2] \subset D'$ car x_0 est point extrême de D' , x_1, x_2 sont par ailleurs d'un même côté de $P(f, \lambda)$ qui est un plan d'appui. Donc deux tels éléments x_1 et x_2 ne peuvent exister et x_0 est point extrême de D .

Théorème 13 : Théorème de Krein-Millmann

La fermeture convexe fermée du profil $\ddot{D}$ d'un convexe compact non vide D est identique à D .

DEMONSTRATION :

Soit $[\ddot{D}]$ la fermeture convexe fermée de $\ddot{D}$

$$\ddot{D} \subset D \Rightarrow [\ddot{D}] \subset [D] = D$$

Il faut maintenant démontrer que $D \subset [\ddot{D}]$

Supposons qu'il existe $x_0 \in D$ tel que $x_0 \notin [\ddot{D}]$. D'après le théorème de séparation, il existerait une forme linéaire f telle que :

$$f(x_0) < \inf [f(x) \mid x \in [\ddot{D}]]$$

Mais $f(D)$ est un compact de $\mathbb{R}$ donc un segment $[\lambda, \mu]$. Le plan d'appui $P(f, \lambda)$ contient un point extrême de D d'après le théorème 12.

Donc $\lambda \in f[\ddot{D}]$. De même, on voit que $\mu \in f[\ddot{D}]$ et comme $f[\ddot{D}]$ est convexe (théorème 5), $f(D) = [\lambda, \mu] \subset f[\ddot{D}]$ d'où la contradiction.

1,3 - POLYEDRES

On a vu, d'après le corollaire du théorème 11, que tout ensemble convexe D est l'intersection de tous les demi-espaces fermés qui le contiennent.

Définition 10 :

Un tronçon D est l'intersection d'un nombre *fini* de demi-espaces fermés.

Par exemple, dans $\mathbb{R}^n$ rapporté à sa base canonique un tronçon T sera de la forme :

$$\sum_j a_{ij} x_j \leq b_i \quad \begin{array}{l} j \in J = [1 \dots n] \\ i \in I = [1 \dots m] \end{array}$$

ou, en écriture matricielle :

$$T = [x \mid Ax \leq b] \quad \begin{array}{l} A \text{ matrice } mn \\ b \text{ matrice } m1 \end{array}$$

Définition 11 :

Un polyèdre convexe est un tronçon borné.

Définition 12 :

Soit x un élément d'un polyèdre D et Vx la variété linéaire formée par les droites privilégiées de D contenant x . Le sous-ensemble $Vx \cap D$ est une *face d'ordre k* si V est de dimension k .

- Une face d'ordre 1 est appelée *arête*. Une face d'ordre 0 correspondant à $Vx = \emptyset$ est un élément extrême ou sommet de D .
- Toute face de D est contenue dans un nombre fini de plans d'appui. D a donc un nombre fini de faces de tous ordres.

Théorème 14 :

Si $X = [x_1 \dots x_p]$ est un ensemble fini de R^n , sa fermeture convexe est un compact.

DEMONSTRATION :

D'après le théorème 6 :

$$[X] = \left\{ \sum_{i=1}^p \lambda_i x_i, \lambda_i \geq 0, \sum_{i=1}^p \lambda_i = 1 \right\}$$

l'application de R^p dans R^n définie par :

$$g(\lambda) = \sum_{i=1}^p \lambda_i x_i, \quad \lambda_i \geq 0, \quad \sum_{i=1}^p \lambda_i = 1$$

est continue car si $\|\lambda\|$ est la norme $\sup [\lambda_i \mid i \in (1 \dots p)]$ et $\|x\|$ une norme dans R^n , à tout $\epsilon > 0$ correspond $\eta = \frac{\epsilon}{\sum \|x_i\|}$ tel que :

$$\|\lambda' - \lambda\| \leq \eta \Rightarrow \|g(\lambda) - g(\lambda')\| = \left\| \sum (\lambda_i - \lambda'_i) x_i \right\| \leq \|\lambda - \lambda'\| \cdot \sum \|x_i\| \leq \epsilon$$

Le sous-ensemble P de R^p tel que $\sum_{i=1}^p \lambda_i = 1, \lambda_i \geq 0$ étant compact $[X] = g(P)$ est compact.

Corollaire 1

Si X est un polyèdre convexe, la fermeture convexe $[\ddot{X}]$ du profil $\ddot{X}$ de X est égale à X .

DEMONSTRATION :

X étant un polyèdre convexe de R^n est compact car fermé et borné. D'après le théorème 13 (Krein Millmann) X est égal à la fermeture convexe fermée du profil $\ddot{X}$ de X . Le profil $\ddot{X}$ étant constitué par des faces d'ordre 0 de X est fini et, d'après le théorème 14, la fermeture $[\ddot{X}]$ de $\ddot{X}$ est compacte, donc fermée. Par suite $X = [\ddot{X}]$.

Ce théorème permet d'obtenir 2 définitions d'un polyèdre :

- 1) Un polyèdre est l'intersection des demi-espaces fermés relatifs à ses plans d'appui. C'est une définition « tangentielle ».
- 2) Un polyèdre est la fermeture convexe de son profil. C'est une définition ponctuelle.

Définition 13 :

Soit D un polyèdre convexe. L'ensemble K tel que $x \in D \Rightarrow \lambda x \in K \quad \lambda \geq 0$ est un cône polyédral convexe.

- Exemple 1

Si le tronçon $Ax \leq b$ est un polyèdre, l'ensemble $Ax \leq 0$ est un cône polyédral convexe.

- Exemple 2

Si $\ddot{X} = [x_1 \dots x_p]$ désigne le profil d'un polyèdre convexe, l'ensemble $K = [x \mid x = \sum_{i=1}^p \lambda_i x_i, \lambda_i \geq 0]$ est un cône polyédral convexe.

1,4 - SYSTEMES D'INEGALITES LINEAIRES

On a vu (théorème 14) qu'un polyèdre pouvait être défini de 2 manières équivalentes, tangentiellement ou ponctuellement. C'est en général le premier mode de définition qui nous intéressera. Ceci implique donc l'étude des principaux théorèmes sur les systèmes d'inégalités linéaires.

Théorème 15 : Théorème de Farkas.

Soit c et x deux éléments de R^n , A une matrice réelle (m, n) une condition nécessaire et suffisante pour que $\langle c, x \rangle \geq 0$ pour tout x tel que $Ax \geq 0$ est qu'il existe $u \geq 0$ tel que $c = {}^t A u$.

DEMONSTRATION :

La condition est suffisante.

Dans R^n , l'ensemble $K = \{x \mid Ax \geq 0\}$ est un cône polyédral convexe:

$$c = {}^t A u, u \geq 0 \Rightarrow \langle c, x \rangle = \langle {}^t A u, x \rangle = \langle u, Ax \rangle \geq 0 \quad \forall x \in K$$

La condition est nécessaire.

a) L'ensemble $K^* = \{y \mid y = {}^t A u, u \geq 0\}$ est un cône appelé cône polaire de K .

$$\left. \begin{array}{l} x \in K \\ y \in K^* \end{array} \right\} \Rightarrow \langle x, y \rangle \geq 0$$

Si $c \notin K^*$ puisque K est fermé, il existe $\bar{y} \in K^*$ tel que :

$$\|c - \bar{y}\| = \inf [\|c - y\|, y \in K^*] > 0$$

b) Soit $v = \bar{y} - c$, on a $\langle v, \bar{y} \rangle = 0$

En effet : la demi-droite $\lambda \bar{y}, \lambda \geq 0$ est contenue dans K^* et la fonction $d(\lambda) = \|c - \lambda \bar{y}\|^2$ doit prendre son minimum pour $\lambda = 1$, donc : $\frac{\langle c - \bar{y}, \bar{y} \rangle}{\|\bar{y}\|^2} = 0$

Par suite : $\langle \bar{y} - c, \bar{y} \rangle = \langle v, \bar{y} \rangle = 0$

c) $y \in K^* \Rightarrow \langle v, y \rangle \geq 0$

En effet : supposons qu'il existe $y' \in K^*$ tel que : $\langle v, y' \rangle < 0$ K étant convexe, tout élément $y_{(\lambda)} = \lambda y' + (1 - \lambda) \bar{y}, \lambda \in [0, 1]$ appartient à K^* .

Le minimum de :

$$d^2_{(\lambda)} = \|c - y_{(\lambda)}\|^2 = \|c - (\lambda y' + (1 - \lambda) \bar{y})\|^2 =$$

$$\|v + \lambda (y' - \bar{y})\|^2 = \|v\|^2 + 2\lambda \langle v, y' - \bar{y} \rangle + \lambda^2 \|y' - \bar{y}\|^2$$

$$\text{s'obtient pour } \lambda = \hat{\lambda} = \frac{-\langle v, y' - \bar{y} \rangle}{\|y' - \bar{y}\|^2} = \frac{\langle v, y' \rangle}{\|y' - \bar{y}\|^2}$$

$\langle v, y' \rangle < 0$ implique $\hat{\lambda} > 0$ en contradiction avec le fait que y correspondant à $\lambda = 0$ minimise $\|c - y_{(\lambda)}\|^2, y \in K^*$.

d) D'après c) $y \in K^* \Rightarrow \langle v, y \rangle \geq 0$ d'où

$$\forall u \geq 0 : \langle v, {}^tAu \rangle = \langle Av, u \rangle \geq 0$$

donc $Av \geq 0$ d'où $v \in K$

$$\text{et } \langle c, v \rangle = \langle \bar{y} - v, v \rangle = -\|v\|^2 < 0$$

$$\text{donc } c \notin K^* \Rightarrow \exists v \in K : \langle cv \rangle < 0$$

ce qui prouve la condition nécessaire.

Corollaire :

Le système ${}^tAu = c, u \geq 0$ admet une solution si, et seulement si, $\langle c, x \rangle \geq 0$ pour tout x tel que $Ax \geq 0$.

CONSEQUENCES DU THEOREME DE FARKAS :

Les conséquences du théorème de Farkas sont extrêmement nombreuses. Nous en développerons trois particulièrement intéressantes.

CONSEQUENCE 1 : s'il existe $\bar{x}$ tel que :

$$|\langle c, \bar{x} \rangle = \inf [\langle c, x \rangle \mid Ax \geq b] | < \infty$$

alors il existe $\bar{y}$ tel que :

$$\langle b, \bar{y} \rangle = \sup [\langle b, y \rangle \mid y \geq 0, {}^tAy = c]$$

$$\text{de plus : } \langle b, \bar{y} \rangle = \langle c, \bar{x} \rangle$$

DEMONSTRATION :

$$\text{Soit } m = \inf [\langle c, x \rangle \mid Ax \geq b]$$

Considérons le système homogène :

$$\begin{aligned} Au - bu_0 &\geq 0 \\ u_0 &\geq 0 \end{aligned}$$

et la conséquence homogène

$${}^tcu - mu_0 \geq 0$$

Si $Ax \geq b$ n'est pas vide, toute solution x sera de la forme

$$x = \frac{u}{u_0} \quad \text{si } u_0 \neq 0$$

$$x = x_0 + \lambda u \quad \lambda \geq 0 \quad \text{si } u_0 = 0$$

Dans ces conditions, d'après le théorème de Farkas,

$$\text{Si } \begin{vmatrix} A - b \\ 0 & 1 \end{vmatrix} \begin{vmatrix} u \\ u_0 \end{vmatrix} \geq 0 \text{ entraine } \begin{vmatrix} {}^tc, -m \end{vmatrix} \begin{vmatrix} u \\ u_0 \end{vmatrix} \geq 0$$

il existe $\begin{vmatrix} \bar{y} \\ \bar{y}_0 \end{vmatrix} \geq 0$ tel que :

$$\begin{vmatrix} {}^tA & 0 \\ -{}^tb & 1 \end{vmatrix} \begin{vmatrix} \bar{y} \\ \bar{y}_0 \end{vmatrix} = \begin{vmatrix} c \\ -m \end{vmatrix}$$

donc il existe $\bar{y} \geq 0$ tel que $c = {}^tA\bar{y}$ et, de plus, ${}^tb\bar{y} - \bar{y}_0 = m$. Puisque $\bar{y}_0 \geq 0$, ${}^tb\bar{y} \geq m$. Le système ${}^tAy = c, y \geq 0$ n'est donc pas vide. Pour tout y satisfaisant à ${}^tAy = c, y \geq 0$ et pour tout x satisfaisant à $Ax \geq b$, on a :

$${}^tcx = {}^tyAx \geq {}^tyb$$

$$\text{donc } m = \inf [\langle cx \rangle \mid Ax \geq b] \geq \sup [\langle b, y \rangle \mid {}^tAy = c, y \geq 0]$$

comme $\langle b\bar{y} \rangle \geq m$, ${}^tA\bar{y} = c$, $\bar{y} \geq 0$:

$$\inf [\langle c, x \rangle \mid Ax = b] = \sup [\langle b, y \rangle \mid {}^tAy = c, y \geq 0]$$

les éléments $\bar{x}$ et $\bar{y}$ sont dits éléments «duaux associés» et les deux problèmes :

- maximiser $\langle c, x \rangle$ sous les conditions $Ax \geq b$
- minimiser $\langle b, y \rangle$ sous les conditions ${}^tAy = c, y \geq 0$

sont dits problèmes duaux.

CONSEQUENCE 2 : on voit de même que si l'un des deux problèmes :

- minimiser $\langle c, x \rangle$ sous les conditions $Ax \geq b, x \geq 0$
- maximiser $\langle b, y \rangle$ sous les conditions ${}^tAy \leq c, y \geq 0$

admet une solution finie, l'autre en admet également une, et que :

$$\inf [\langle cx \rangle \mid Ax \geq b, x \geq 0] = \sup [\langle by \rangle \mid {}^tAy \leq c, y \geq 0]$$

CONSEQUENCE 3 :

a) S'il existe x tel que $Ax \geq b, x \geq 0$, alors

$$\inf [\langle c, x \rangle \mid Ax \geq b, x \geq 0] = -\infty$$

si, et seulement si, le système ${}^tAy \leq c, y \geq 0$ est vide.

b) S'il existe y tel que ${}^tAy \leq c, y \geq 0$, alors :

$$\sup [\langle b, y \rangle \mid {}^tAy \leq c, y \geq 0] = +\infty$$

si, et seulement si, le système $Ax \geq b, x \geq 0$ est vide.

DEMONSTRATION de 3 a) :

$\inf [\langle c, x \rangle \mid Ax \geq b, x \geq 0] = -\infty$ implique que le système $Ax \geq b, x \geq 0, {}^tcx \leq \lambda$ n'est pas vide pour tout λ et réciproquement

Il existe donc $v_1 = (v_1 \ v_2 \ v_3) \geq 0$ tel que :

$$A v_1 - v_2 = b$$

$${}^t c v_1 + v_3 = \lambda$$

D'après le théorème de Farkas, ce dernier système ne sera pas vide si, et seulement si, $-{}^t b u_1 + \lambda u_2 \geq 0$,

pour tout u_1, u_2 tel que

$${}^t A u_1 \leq c u_2$$

$$u_1 \geq 0$$

$$u_2 \geq 0$$

Ceci devant être vrai pour tout λ implique $u_2 = 0$. Donc, il ne peut y avoir de solution au système ${}^t A y \leq c, y \geq 0$. On démontrerait de même la conséquence 3 b).

Théorème 16 :

Si $\bar{x}$ est solution du problème :

$$\sup [\langle c, x \rangle \mid A x \leq b, x \geq 0] \quad (1)$$

et $\bar{y}$ solution du problème :

$$\inf [\langle b, y \rangle \mid {}^t A \bar{y} \geq c, y \geq 0] \quad (2)$$

$$\text{alors : } {}^t \bar{y} (A \bar{x} - b) = 0$$

$${}^t \bar{x} ({}^t A \bar{y} - c) = 0$$

Réciproquement, si $\bar{x}$ et $\bar{y}$ vérifient ces 2 dernières relations et si :

$$\bar{x} \in U = [x \mid A x \leq b, x \geq 0]$$

$$\bar{y} \in V = [y \mid {}^t A y \geq c, y \geq 0]$$

x et y sont solutions des problèmes 1 et 2,

DEMONSTRATION :

$\bar{x}$ et $\bar{y}$ solutions de 1 et 2 respectivement impliquent :

$$0 = {}^t c \bar{x} - {}^t b \bar{y} \leq {}^t \bar{y} {}^t A \bar{x} - {}^t \bar{y} b = {}^t \bar{y} (A \bar{x} - b)$$

$$\text{or : } {}^t \bar{y} \geq 0 \quad A \bar{x} \leq b \quad \text{d'où } {}^t \bar{y} (A \bar{x} - b) = 0$$

$$\text{On obtient de même : } {}^t \bar{x} ({}^t A \bar{y} - c) = 0$$

Réciproquement, si $\bar{x} \in U, \bar{y} \in V$ et si $\bar{x}$ et $\bar{y}$ vérifient

$${}^t \bar{y} (A \bar{x} - b) = 0 \quad {}^t \bar{x} ({}^t A \bar{y} - c) = 0$$

alors ${}^t c \bar{x} = {}^t b \bar{y}$, donc $\bar{x}$ et $\bar{y}$ sont solutions des problèmes 1 et 2 respectivement

REMARQUES :

- 1) Le théorème de Farkas peut se déduire aussi du théorème de séparation (théorème 11).
- 2) Les relations du théorème 16 ${}^t\bar{y} (A\bar{x} - b) = 0$ ${}^t\bar{x} ({}^tA\bar{y} - c) = 0$ appelées conditions d'exclusions peuvent s'obtenir plus généralement à partir du théorème de Kuhn et Tucker pour les problèmes concaves.
- 3) Il est équivalent de résoudre le problème maximiser $\langle c, x \rangle$ compte tenu de $Ax \leq b$ $x \geq 0$ ou le problème trouver que x et y tel que :

$$\begin{array}{ll} Ax \leq b & x \geq 0 \\ {}^tAy \geq c & y \geq 0 \\ {}^tcx \geq {}^tb y \end{array}$$

LA METHODE DITE DU SIMPLEXE

2,1 - GENERALITES

2,11 - POSITION DU PROBLEME

Un programme linéaire se pose en toute généralité sous la forme canonique suivante :

$$\text{maximer } \sum_{j \in N} c_j x_j$$

compte tenu de

$$\sum_i a_{ij} x_j \leq b_i \quad i \in I = [1 \dots m]$$

$$x_j \geq 0 \quad j \in N = [1 \dots n]$$

a_{ij}, b_i, c_j étant des valeurs réelles.

La forme linéaire $\sum_{j \in N} c_j x_j$ est fréquemment appelée fonction économique et les relations $\sum_{j \in N} a_{ij} x_j \leq b_i$ sont appelées contraintes.

Un programme linéaire peut s'écrire sous forme condensée (3) $\sup [\langle c, x \rangle \mid Ax \leq b, x \geq 0]$ A étant une matrice réelle m, n .

Le sous-ensemble D de R^n

$$D = [x \mid Ax \leq b, x \geq 0]$$

est appelé le domaine *des solutions admissibles*. C'est par définition (définitions 10 et 11) un polyèdre de R^n si D est borné, et on peut toujours se ramener à ce cas en ajoutant, si besoin est, la contrainte $\sum_{j \in N} x_j \leq M$ M étant une constante positive arbitrairement grande

2,12 - FORMES EQUIVALENTES

Le problème 3 est de façon évidente équivalent au problème suivant :

$$(4) \sup [\langle c, x \rangle \mid Ax + y = b, x \geq 0, y \geq 0], x \in R^n, y \in R^m$$

lui-même équivalent au problème suivant :

$$\sup [\langle l, z \rangle \mid z = x + y, Ax + y = b, z \geq 0], l = (c, 0) \in R^{m+n}$$

Le domaine D' de R^{m+n}

$$D' = [z \mid z = x + y, Ax + y = b, x \geq 0, y \geq 0]$$

est par définition un tronçon dans R^{m+n}

Soit f l'application linéaire de R^{m+n} dans R^m de matrice (A, I) par rapport aux bases canoniques $\epsilon_1, \dots, \epsilon_{n+m}$ de R^{n+m} et $e_1, \dots, e_m$ de R^m .

Il est immédiat que :

$$D' = f^{-1}(b) \cap K$$

où $f^{-1}(b)$ est l'image réciproque de $b \in R^m$ dans R^{m+n} , et K le cône polyédral convexe engendré par les vecteurs ϵ_i de la base canonique de R^{m+n}

$f^{-1}(b)$ est un sous-espace affine de R^{m+n} parallèle au noyau N de l'application f donc de même dimension que N . Comme l'application f est de rang m , N sera de dimension n . D' a par suite un intérieur vide dans R^{m+n}

S'il existe $z \in D'$ tel que $z_i > 0 \quad \forall i \in (1 \dots m+n)$, D' est de dimension n .

Théorème 17 : D et D' sont isomorphes

DEMONSTRATION :

Soit : $z = x + y \in D'$ et $x \in D$, l'application $z \mapsto x$ est visiblement une projection, donc surjective et linéaire. Elle est aussi injective car $x_1 = x_2$ entraîne $y_1 = y_2$ ($= b - Ax_1 = b - Ax_2$) donc $z_1 = z_2$ p est donc bijective, par suite D et D' sont isomorphes en particulier à tout point extrême de D' correspond un point extrême de D et réciproquement.

Théorème 18 :

A tout sous-ensemble de m indices $J_0 \subset J = [1 \dots m+n]$ tel que :

1. La famille $f(\epsilon_j)$, $j \in J_0$ soit libre dans R^m
2. $\sum_{j \in J_0} \lambda_j f(\epsilon_j) = b \quad \lambda_j \geq 0$

correspond un point extrême z de D' tel que :

$$z = [z_j = \lambda_j, j \in J_0 ; z_j = 0, j \in J - J_0]$$

DEMONSTRATION :

Il est immédiat que tout z ainsi défini appartient à D' . Si un élément z ainsi défini n'était pas point extrême de D' , il existerait une droite privilégiée δ et sur cette droite 2 éléments z^1 et $z^2 \in D'$ tels que :

$$z = \frac{z^1 + z^2}{2}$$

$$\text{or } \forall j \in J - J_0, z_j = 0 \Rightarrow z^1_j = 0, z^2_j = 0$$

Donc z^1 et z^2 sont tels que :

$$\sum_{j \in J_0} z^1_j f(\epsilon_j) = b \quad z^1_j \geq 0$$

$$\sum_{j \in J_0} z^2_j f(\epsilon_j) = b \quad z^2_j \geq 0$$

donc : $\sum_{j \in J_0} (z_j^1 - z_j^2) f(\epsilon_j) = 0$ ce qui est impossible

puisque la famille $f(\epsilon_j)_{j \in J_0}$ est libre dans R^m .

Théorème 19 :

A tout point extrême de D' correspond $J_0 \subset J$ tel que :

1. $|J_0| \leq m$
2. $f(\epsilon_j)_{j \in J_0}$ est une famille libre de R^m .

DEMONSTRATION :

Soit : z un point extrême de D' tel que :

$$z = [z_j = \lambda_j \quad j \in J_0 ; z_j = 0 \quad j \in J - J_0]$$

$$\sum_{j \in J_0} \lambda_j f(\epsilon_j) = b \quad \lambda_j > 0$$

Supposons que la famille $f(\epsilon_j)_{j \in J_0}$ ne soit pas libre dans R^m on aurait $\sum_{j \in J_0} \mu_j f(\epsilon_j) = 0$ avec

$\mu_j \neq 0$ pour au moins un j . Par suite, il existerait $u \in R^{m+n}$ tel que :

$$u = [u_j = \mu_j, \quad j \in J_0 ; u_j = 0 \quad j \in J - J_0]$$

et comme $z_j > 0 \quad \forall j \in J_0$, il existerait $\theta \neq 0$ tel que :

$$z^1 = z + \theta u \in D'$$

$$z^2 = z - \theta u \in D'$$

d'où $z = \frac{z^1 + z^2}{2}$ ne serait pas point extrême de D' .

Théorème 20 :

Si le problème $\sup \{ \langle 1, z \rangle \mid z \in D' \}$ a une solution, il existe au moins un point extrême $\hat{z}$ de D' tel que :

$$\langle 1, \hat{z} \rangle = \sup \{ \langle 1, z \rangle \mid z \in D' \}$$

DEMONSTRATION :

soit : $\bar{z} \in D'$ tel que $\langle 1, \bar{z} \rangle = \sup \{ \langle 1, z \rangle \mid z \in D' \}$ et $\bar{D}' = [z_1 \dots z_p]$ le profil de D' .

$$\bar{z} \in D' \Rightarrow \bar{z} = \sum_{i=1}^p \lambda_i z_i \quad \lambda_i \geq 0 \quad \sum_{i=1}^p \lambda_i = 1$$

$$\text{donc } \langle 1, \bar{z} \rangle = \sum_{i=1}^p \lambda_i \langle 1, z_i \rangle \leq \sup_{i=1}^p \langle 1, z_i \rangle \sum_{i=1}^p \lambda_i = \sup_{i=1}^p \langle 1, z_i \rangle = \langle 1, \hat{z} \rangle$$

comme : $\hat{z} \in D' \quad \langle 1, \hat{z} \rangle \leq \langle 1, \bar{z} \rangle$

donc : $\langle 1, \hat{z} \rangle = \langle 1, \bar{z} \rangle$

2,2 - LA METHODE SIMPLEXE

2,21 - LA METHODE DIRECTE

S'il existe des solutions au problème posé, d'après le théorème 20, ces solutions sont à chercher parmi les éléments du profil $\bar{D}'$ de D' . D'après les théorèmes 18 et 19, tout élément de $\bar{D}'$ sera déterminé par un sous-ensemble de m indices $J_0 \subset J$ tels que :

1. La famille $f(\epsilon_j) \quad j \in J_0$ forme une base de R^m .
2. b ait des composants positives ou nulles dans cette base.

REMARQUE :

- . Par abus de langage on dira que $J_0 \subset J$ forme une base admissible pour le problème, si les conditions 1 et 2 précédentes sont vérifiées.
- . Une base admissible telle que l'élément extrême correspondant maxime $\langle 1, z \rangle$ dans D sera dite optimale.

2,211 - PASSAGE D'UN POINT EXTREME A UN AUTRE

Supposons connue une base admissible $J_0 \subset J$ donc un élément $z^0 \in \bar{D}'$ tel que :

$$z^0 = [z_j^0 \geq 0 \quad j \in J_0 \quad ; \quad z_j^0 = 0 \quad j \in J - J_0]$$

on aura :

$$\sum_j z_j^0 f(\epsilon_j) = b, \quad j \in J_0$$

$$\langle 1, z^0 \rangle = \sum_j 1_j z_j^0$$

soit $k \notin J_0$ et $\theta \geq 0$ on aura encore :

$$\sum_j z_j^0 f(\epsilon_j) + \theta f(\epsilon_k) - \theta f(\epsilon_k) = b$$

$$k \notin J_0 \text{ entraîne } f(\epsilon_k) = \sum_{j \in J_0} z_{jk} f(\epsilon_j)$$

par suite :

$$\sum_{j \in J_0} (z_j^0 + \theta z_{jk}) f(\epsilon_j) + \theta f(\epsilon_k) = b$$

si :

$$0 \leq \theta \leq \inf_{j \in J_k^+} \left(\frac{z_j^0}{z_{jk}} \right) = \frac{z_r^0}{z_{rk}}, \quad J_k^+ = [j \in J_0 \mid z_{jk} > 0]$$

l'élément correspondant

$$z(\theta) = [z_j^0 - \theta z_{jk}, j \in J_0; \theta, j = k; 0, j \in J - J_0 - [k]]$$

appartient à D' .

La variation de la forme $\langle l, z \rangle$ sera :

$$\langle l, z(\theta) \rangle - \langle l, z^0 \rangle = \theta \left(l_k - \sum_{j \in J_0} l_j z_{jk} \right) = \theta (l_k - p_k) = -\theta d_k$$

$$\text{avec : } p_k = \sum_{j \in J_0} l_j z_{jk} \quad d_k = p_k - l_k$$

si $d_k < 0$ le maximum en θ de $\langle l, z(\theta) \rangle$ sera obtenu pour $\theta = \frac{z_r^0}{z_{rk}}$ qui correspond à un nouveau point extrême z^1 si $z_r^0 \neq 0$ car : $J_1 = J_0 + [k] - [r]$, constitue une nouvelle base admissible.

On voit aisément que dans la nouvelle base J_1 :

. les coordonnées de b seront :

$$z_j^1 = z_j^0 - z_r^0 \frac{z_{jk}}{z_{rk}} \quad j \in J_1 - [k]$$

$$z_k^1 = \frac{z_r^0}{z_{rk}}$$

. les composantes de $f(\epsilon_i)$, $i \in J$ seront :

$$z'_{ji} = z_{ji} - \frac{z_{ri}}{z_{rk}} z_{jk} \quad j \in J_1 - [k]$$

$$z'_{ki} = \frac{z_{ri}}{z_{rk}}$$

Il en résulte d'après la définition des $d_i = l_i - \sum_{j \in J_0} l_j z_{ji}$ que : $d'_i = d_i - d_k \frac{x_{ri}}{x_{rk}}$ et

$$\text{la nouvelle fonction économique sera : } \langle l, z^1 \rangle = \langle l, z^0 \rangle - \frac{z_r^0}{z_{rk}} d_k$$

2,212 - THEOREME 21 :

Si $z^0 \in D'$ il existe 3 éventualités exclusives et exhaustives.

1. Il existe $k \in J$ tel que $d_k < 0$ et pour tout $j \in J_0, z_{jk} \leq 0$. Dans ces conditions $z(\theta)$

appartient à D' pour tout $\theta > 0$ et $\langle 1, z(\theta) \rangle$ peut être rendue arbitrairement grande. Le maximum est infini.

2. Pour tout $k \in J$ tel que $d_k < 0$, il existe $j(k) \in J_0$ tel que $z_{j(k),k} > 0$. Dans ce cas, on peut trouver un autre point extrême z^1 tel que $\langle 1, z^1 \rangle \geq \langle 1, z^0 \rangle$
3. Pour tout $k \in J, d_k \geq 0$. Le maximum est atteint en z^0 . La démonstration de 1 et 2 est évidente.

Démonstration de 3 :

En effet, soit $z \in D', z \neq z^0$

on a : $f(z) = b \quad z \geq 0$

donc : $\sum_{k \in J} z_k f(\epsilon_k) = b \quad z_k \geq 0$

$\forall k \in J, d_k \geq 0 \Rightarrow p_k \geq l_k$ donc :

$$\langle 1, z \rangle = \sum_{k \in J} l_k z_k \leq \sum_{k \in J} p_k z_k = \sum_{k \in J} \left(\sum_{j \in J_0} l_j z_{jk} \right) z_k = \sum_{j \in J_0} \left(\sum_{k \in J} z_{jk} z_k \right) l_j$$

or comme : $f(z^0) = b$
 $f(z) = b$

$$b = \sum_{j \in J_0} z_j^0 f(\epsilon_j) = \sum_{k \in J} z_k f(\epsilon_k) = \sum_{j \in J_0} \left(\sum_{k \in J} z_k z_{jk} \right) f(\epsilon_j)$$

J_0 étant une base de R^m , b admet une décomposition unique donc :

$$z_j^0 = \sum_{k \in J} z_{jk} z_k$$

et par suite :

$$\langle 1, z \rangle \leq \sum_{j \in J_0} l_j z_j^0 = \langle 1, z^0 \rangle$$

Comme il existe un nombre fini d'éléments extrêmes et qu'un même élément extrême ne peut être formé 2 fois puisque $\langle 1, z \rangle$ est croissante. Le maximum sera atteint en un nombre fini d'étapes.

REMARQUE :

Les cas où $z_r = 0$ seront appelés cas de dégénérescence, ils conduisent à $\langle 1, z \rangle$ stationnaire. Une procédure a été mise au point par Charnes pour s'en affranchir en un nombre fini d'étapes.

2.22 - LA METHODE DUALE

Définition 14 :

On appellera base duale admissible une base $J_0 \subset J$ telle que :

1. La famille $f(\epsilon_j), j \in J_0$ soit libre dans R^m .
2. Les coefficients $d_j, j \in J$ soient tous non négatifs.

Pour toute base duale admissible $b \in \mathbb{R}^m$ admet une décomposition telle que :

$$\sum_{j \in J_0} \lambda_j f(\epsilon_j) = b$$

l'élément $z^0 \in \mathbb{R}^{m+n}$ tel que :

$$z^0 = [z_j^0 = \lambda_j, j \in J_0; z_j = 0, j \in J - J_0]$$

sera dit *point pseudo-extrême*.

PROPRIETES :

1. En tout point pseudo-extrême $z^0 < 1, z^0 > \geq < 1, z > z \in D'$. (Ceci résulte immédiatement de ce qui a été vu au point 3 du paragraphe précédent (la démonstration n'utilise que l'hypothèse $f(z^0) = b$).)
2. Il existe un nombre fini de points pseudo-extrêmes. En effet, il existe au plus $C \frac{m}{m+n}$ bases duales admissibles...

Théorème 21 bis :

- a. Si les composantes $z_j^0, j \in J_0$ sont non négatives l'élément z^0 maximise $< 1, z >$ dans D' .
- b. S'il existe $r \in J_0$ tel que $z_r^0 < 0$ et que pour tout $i \in J$ les composantes z_{ri} de tout $f(\epsilon_i)$ dans la base J_0 de $\mathbb{R}^m$ sont positives; D' est vide.
- c. Si pour tout $r \in J_0$ tel que $z_r^0 < 0$ il existe au moins une composante z_{rj} de $f(\epsilon_j)$ dans la base J_0 de $\mathbb{R}^m$, négative, on peut trouver un autre point pseudo-extrême z^1 tel que: $< 1, z^1 > \leq < 1, z^0 >$.

DEMONSTRATION :

a. est évident d'après le théorème 21.

b. soit $z \in D'$ donc tel que : $f(z) = b, z \geq 0$, équivalent à :

$$\sum_{i \in J} z_i f(\epsilon_i) = b \quad z_i \geq 0 \quad \forall i \in J$$

J_0 constituant une base de $\mathbb{R}^m$ on a donc :

$$f(\epsilon_i) = \sum_{j \in J_0} z_{ji} f(\epsilon_j)$$

puis :

$$b = \sum_{j \in J_0} \left(\sum_{i \in J} z_{ji} z_i \right) f(\epsilon_j)$$

Par ailleurs, z^0 étant point pseudo-extrême est tel que $f(z^0) = b$ donc tel que :

$$\sum_{j \in J_0} z_j^0 f(\epsilon_j) = b$$

L'expression de b devant être unique dans la base J_0 on a :

$$\sum_{i \in J} z_{ji} z_i = z_j^0 \quad \forall j \in J_0$$

Il est impossible de trouver des $z_i \geq 0$ tels que la relation

$$\sum_{i \in J} z_{ri} z_i = z_r^0$$

soit vérifiée si $z_{ri} \geq 0 \quad \forall i$ et $z_r < 0$

Donc $D' = \emptyset$

c. Supposons que pour tout $r \in J_0$ tel que $z_r^0 < 0$ il existe au moins une composante $z_{rj} < 0$.
S'il existe un point pseudo-extrême z^1 voisin de z^0 et tel que $\langle 1, z^1 \rangle \leq \langle 1, z^0 \rangle$, il existe $r \in J_0$ et $k \in J - J_0$ tels que :

$$(1) \quad \frac{-z_r^0}{z_{rk}} d_k = \langle 1, z^1 \rangle - \langle 1, z^0 \rangle \leq 0$$

la base $J_1 = J_0 + [k] - [r]$ doit être duale admissible donc :

$$(2) \quad d'_j = d_j - d_k \frac{z_{rj}}{z_{rk}} \geq 0 \quad \forall j \in J$$

$$(1) \text{ implique : } \frac{z_r^0}{z_{rk}} \geq 0$$

la relation (2) et le fait que $k \neq r$ impliquent $z_{rk} < 0$

donc $z_r^0 \leq 0$ et :

$$\frac{d_k}{z_{rk}} = \inf_{j \in J_r^-} \left(\frac{d_j}{z_{rj}} \right) \text{ avec } J_r^- = [j \mid j \in J, z_{rj} < 0]$$

D'après l'hypothèse de tels indices r et k existent. Donc on peut trouver z^1 tel que :

$$\langle 1, z^1 \rangle \leq \langle 1, z^0 \rangle.$$

Comme il existe un nombre fini de points pseudo-extrêmes et que sauf cas de dégénérescence pour lesquels une procédure a été mise au point par Charnes $\langle 1, z \rangle$ est strictement décroissante, la solution du problème sera obtenue en un nombre fini d'étapes.

2.23 - INTERPRETATION DE LA METHODE SIMPLEXE EN TERME DE GEOMETRIE AFFINE : LA MATRICE SIMPLEXE.

Pour résoudre les programmes linéaires, il est commode d'utiliser la matrice simplexe qui représente les vecteurs $b, f(\epsilon_j) \quad j \in J$ dans une base $J_0 \subset J$ de $\mathbb{R}^m$ et d'ajouter à cette matrice la ligne des $d_j \quad j \in J$.

La matrice simplexe canonique correspond au cas où la base de $\mathbb{R}^m$ est la base canonique.

Le programme linéaire $\sup [\langle c, x \rangle \mid Ax \leq b, x \geq 0]$ équivaut au programme :

$$\sup [\langle l, z \rangle \mid Ax + y = b, z = x + y, z \geq 0, l = (c, 0)]$$

aura comme matrice canonique :

$$S_0 = \begin{pmatrix} b & A & I \\ 0 & -c & 0 \end{pmatrix}$$

On sait qu'à toute base admissible de R^m correspond un point extrême de D' .

En généralisant cette notion :

Définition 15 :

A tout ensemble de m indices $J_0 \subset J = [1, \dots, m+n]$ tels que les $f(\epsilon_j)$ $j \in J_0$ soient libres dans R^m correspond l'élément $z(J_0)$ de R^{m+n} appelé point extrême généralisé de D tel que :

$$\text{si } b = \sum_{j \in J_0} \lambda_j f(\epsilon_j)$$

$$z(J_0) = [z_j = \lambda_j, j \in J_0; z_j = 0, j \in J - J_0]$$

Puisque la matrice représentant les $f(\epsilon_j)$ $j \in J$ dans la base canonique de R^m est (A, I) , $z(J_0)$ vérifie les relations :

$$\sum_{j \in N} a_{ij} z_j + z_{n+i} = b_i \quad \begin{array}{l} i \in I = [1, \dots, m] \\ j \in N = [1, \dots, n] \end{array}$$

Partition de la matrice simplexe :

Soit $\bar{J}_0 = J - J_0$, $\bar{J}_0$ comporte n indices.

Puisque $J = [1, \dots, m+n]$ on peut écrire :

$$\text{d'une part : } J = \bar{J}_0 \cup J_0 \quad \bar{J}_0 \cap J_0 = \emptyset$$

$$\text{d'autre part } J = N \cup (n+I) \quad N \cap (n+I) = \emptyset$$

ce qui conduit à la partition de J

$$\begin{aligned} J &= (J_0 \cup \bar{J}_0) \cap (N \cup (n+I)) = (J_0 \cap N) \cup (\bar{J}_0 \cap N) \cup (\bar{J}_0 \cap (n+I)) \cup (J_0 \cap (n+I)) \\ &= J_1 \cup J_2 \cup J_3 \cup J_4 \end{aligned}$$

$$\text{Puis comme } I = (n+I) - n \quad \text{et} \quad n+I = J_3 \cup J_4$$

$$\text{On a : } I = (J_3 - n) \cup (J_4 - n)$$

$$\text{ou } I = I_1 \cup I_2$$

Il en résulte la partition suivante de la matrice canonique :

		J_1	J_2	J_3	J_4
I_1 $S_0 \equiv I_2$	b_1	A_1	A_3	I	o
	b_2	A_2	A_4	o	I
	o	c_1	c_2	o	o

Si R^m est rapporté à la base $f(\epsilon_j)$, $j \in J_0 = J_1 \cup J_4$ la matrice du changement de base est :

$$\begin{vmatrix} A_1 & o \\ A_2 & I \end{vmatrix}$$

elle est évidemment régulière puisque la famille $f(\epsilon_j)$ $j \in J_0$ est libre dans R^m .

Son inverse est :

$$\begin{vmatrix} A^{-1} & o \\ -A_2 A_1^{-1} & I \end{vmatrix}$$

Par suite, la matrice simplexe correspondante sera :

		J_1	J_2	J_3	J_4
J_1 $S(J_0)$	J_1	$A_1^{-1}b$	I	$A_1^{-1}A_3$	A_1^{-1}
	J_4	$b_2 - A_2 A_1^{-1}b_1$	o	$A_4 - A_2 A_1^{-1}A_3$	$-A_2 A_1^{-1}$
		$c_1 A_1^{-1}b_1$	o	$c_1 A_1^{-1}A_3 - c_2$	$c_1 A_1^{-1}$

avec $J_1 \cup J_4 = J_0$, $J_2 \cup J_3 = \bar{J}_0$

Signification en termes de géométrie affine

Faisons correspondre à tout $j \in [0 \dots n+m] = (o) \cup J$ les applications f_j de R^n dans R telles que :

$$f_j(x) = u_j + \langle v_j, x \rangle$$

. u_j étant un scalaire égal à

$$\begin{vmatrix} o & \text{pour } j = o \\ o & \text{pour } j \in N \\ b_i & \text{pour } j-n \in I \end{vmatrix}$$

. v_j étant un vecteur de R^n représenté dans la base canonique de R^n par :

$$\begin{array}{ll} (c_1 \dots c_n) & \text{si } j = 0 \\ (0 \dots 1 \dots 0) & \text{si } j \in N \\ (-a_{j1}, \dots, -a_{jn}) & \text{si } j-n \in I \end{array}$$

Le sous ensemble de $R^n : [x \mid f_j(x) = 0]$ est un plan désigné par P_j .

Le sous ensemble de $R^n : [x \mid f_j(x) \geq 0]$ est un demi-espace fermé désigné par H_j .

Définition 16

Si H_j est le demi-espace $[x \mid u_j + \langle v_j, x \rangle \geq 0]$ on appellera normale intérieure relative au plan P_j défini par $[x \mid u_j + \langle v_j, x \rangle = 0]$ le vecteur v_j de R^n .

La justification de cette définition provient du fait que si $x_0 \in P_j$ l'élément $x_1 = x_0 + \lambda v_j$, $\lambda \geq 0$ appartient au demi-espace fermé H_j .

Considérons la matrice M_0 à $m + n + 1$ lignes et $n + 1$ colonnes définie par :

$$M_0 = (U, V) = \begin{vmatrix} 0 & c \\ 0 & I \\ b & -A \end{vmatrix}$$

La matrice U représente la position en grandeur et en signe de l'origine par rapport aux plans P_j , $j \in (0) \cup J$ et les lignes de la matrice V représentent dans la base canonique de R^n les composantes des normales intérieures aux plans P_j , $j \in (0) \cup J$.

Soit maintenant $q_1(x) = p_1 + \langle q_1, x \rangle$, $1 \in L$, n applications de R^n dans R telles que les n vecteurs q_1 soient linéairement indépendants.

Soit q la matrice dont les lignes représentent les vecteurs q_1 dans la base canonique de R^n et p la matrice $(n, 1)$ formée des scalaires p_1 .

Considérons le cône affine K_L de R^n engendré par les demi-espaces : $q_1(x) \geq 0$ $1 \in L$

On aura $K_L = [x \mid p + qx \geq 0]$

ou : $K = [x \mid q(x - x_L) \geq 0]$ en appelant x_L le sommet du cône K_L , (soit $x_L = -q^{-1}p$).

La position de x_L par rapport aux plans P_j , $j \in (0) \cup J$ est décrite par la matrice U' .

$$U' = U - Vq^{-1}p$$

Les composantes des normales intérieures aux plans P_j , $j \in (0) \cup J$ dans la base des normales intérieures aux plans P_1 sont données par : $U' = Vq^{-1}$ de telle sorte que la matrice M_0 est devenue :

$$M_0 = M_0' \cdot \begin{vmatrix} 1 & 0 \\ -q^{-1}p & q^{-1} \end{vmatrix} = M_0' U^{-1}$$

avec $N = \begin{vmatrix} 1 & 0 \\ p & q \end{vmatrix}$

CAS PARTICULIER :

En introduisant la partition précédente de $J = J_1 \cup J_2 \cup J_3 \cup J_4$ on aura :

$$M_O = \begin{array}{c|ccc} & & J_1 & J_2 \\ \hline o & o & c_1 & c_2 \\ J_1 & o & I & o \\ J_2 & o & o & I \\ J_3 & b_1 & -A_1 & -A_3 \\ J_4 & b_2 & -A_2 & -A_4 \end{array}$$

Si l'on prend pour ensemble L l'ensemble $\bar{J}o = J_2 \cup J_3$ on aura :

$$p = \begin{vmatrix} o \\ b_1 \end{vmatrix} \quad q = \begin{vmatrix} o & I \\ -A_1 & -A_3 \end{vmatrix}$$

d'où :

$$q^{-1} = \begin{vmatrix} -A_1^{-1}A_3 & -A_1^{-1} \\ I & o \end{vmatrix}$$

puis :

$$N^{-1} = \begin{vmatrix} I & o \\ -q^{-1}p & q^{-1} \end{vmatrix}$$

et enfin :

$$M(\bar{J}o) = \begin{array}{c|ccc} & o & J_2 & J_3 \\ \hline o & c_1 A_1^{-1} b_1 & c_2 - c_1 A_1^{-1} A_3 & -c_1 A_1^{-1} \\ J_1 & A_1^{-1} b_1 & -A_1^{-1} A_3 & -A_1^{-1} \\ J_2 & o & I & o \\ J_3 & o & o & I \\ J_4 & b_2 - A_2 A_1^{-1} b_1 & -A_4 + A_2 A_1^{-1} A_3 & A_2 A_1^{-1} \end{array}$$

On énoncera donc :

Théorème 22

Si $z(Jo)$ est un point extrême généralisé de D' les normales intérieures aux plans P_j , $j \in \bar{J}o$ forment une base de R^n .

Puis par comparaison de $M(\bar{J}_0)$ avec la matrice simplexe $S(J_0)$ on énoncera :

Théorèmes 23 :

Si $z(J_0)$ est un point extrême généralisé de D' , les plans $P_j, j \in \bar{J}_0 = J - J_0$ admettent une intersection réduite à un seul élément $x(\bar{J}_0)$ dont la position par rapport aux plans $P_j, j \in J_0$, est donnée par la première colonne de la matrice simplexe et la restriction de la matrice simplexe à ses colonnes d'indice dans $\bar{J}_0$ représente les composantes des normales *extérieures* aux plans $P_j, j \in J_0 \cup \{0\}$ dans la base de R^n formée par les normales *intérieures* aux plans $P_j, j \in \bar{J}_0$.

L'élément $x(\bar{J}_0)$ sera dit point extrême généralisé du polyèdre D de R^n .

REMARQUES :

1. Le cône $K(\bar{J}_0)$ engendré par les demi-espaces $H_j, j \in \bar{J}_0$ de sommets $x(J_0)$ contient $D = \bigcap_{j \in J} H_j$

La méthode simplexe directe consiste à définir des cônes $k(\bar{J}_0)$ dont le sommet est situé sur D et à augmenter à chaque étape la forme $\langle c, x(\bar{J}_0) \rangle$. Quand un cône $K(\bar{J}_0)$ est tel que l'opposé du gradient c de $\langle c, x \rangle$ appartient au cône $K^*(\bar{J}_0)$ formé par les normales intérieures aux plans $P_j, j \in \bar{J}_0$ alors, d'après le théorème de Farkas $\langle c, x \rangle$ est maximum dans $K(\bar{J}_0)$ donc dans D puisque $x(\bar{J}_0) \in D$ et $D \subset K(\bar{J}_0)$.

2. La méthode duale de la méthode simplexe consiste à cheminer de points pseudo-extrêmes en points pseudo-extrêmes; $x(\bar{J}_0)$ maximise $\langle c, x \rangle$ dans $K(\bar{J}_0)$ et quand $x(\bar{J}_0) \in D$ le maximum de $\langle c, x \rangle$ dans D est atteint.
3. Il est possible de définir une méthode mixte de résolution des programmes linéaires : partant d'un point extrême généralisé quelconque x_0 , on maximisera par la méthode directe la forme $\langle c, x \rangle$ dans le polyèdre D'' .

$$D'' = \bigcap_{l \in L} H_l$$

$$L = [j \mid j \in J \mid u_j + \langle V_j, x_0 \rangle \geq 0]$$

puis, ce maximum étant atteint, ou bien le point extrême $\hat{x}$ de D'' correspondant appartient à D et le problème est résolu, ou bien il n'appartient pas à D mais $\hat{x}$ est point pseudo-extrême pour D et, par la méthode duale de la méthode simplexe, on formera la solution du problème.

Il est évidemment possible aussi de procéder en sens inverse, c'est-à-dire ayant défini x_0 point extrême généralisé de D de restreindre le problème au sous-espace V de R^n ou x_0 est point pseudo-extrême de $V \cap D$ puis, par la méthode duale de trouver le maximum de $\langle c, x \rangle$ dans $D \cap V$. Si l'élément x correspondant dans D est aussi pseudo-extrême c'est l'optimum du problème posé sinon, on continuera par la méthode simplexe directe. Cette méthode mixte à l'avantage suivant : il est souvent possible dans un problème pratique d'envisager a priori quelles équations seront strictement satisfaites et quelles variables seront nulles. De cette manière, le point extrême généralisé correspondant constitue soit l'optimum du problème soit une approximation assez voisine de la solution qu'on obtiendra par la méthode mixte.

4. Si $x(\bar{J}_0)$ est point extrême généralisé de D , défini dans R^n par $D = [x \mid Ax \leq b, x \geq 0]$ alors l'élément dual associé $u(J_0)$ de $x(\bar{J}_0)$ est point extrême généralisé du polyèdre D^* défini par $D^* = [u \mid uA \geq c, u \geq 0]$ et contenu dans R^m , considéré comme son propre dual ($R^{m*} = R^m$).

5. Si x est point extrême de D , u est point pseudo-extrême de D^* et réciproquement.

D et D^* ont même nombre de points extrêmes généralisés, il y en a au plus C_{m+n}^n et le nombre de points extrêmes de D (resp. de D^*) est égal au nombre de points pseudo-extrêmes de D^* (resp. de D).

Ecrivons les contraintes :

$$\left| \begin{array}{l} \sum_{j \in N} a_{ij} x_j \leq b_i \\ x_j \geq 0 \end{array} \right. \quad \begin{array}{l} i \in I = [1 \dots m] \\ j \in N = [1 \dots n] \end{array}$$

Sous la forme symétrique $\langle q_i, x \rangle + p_i > 0 \quad \forall i \in J = (N \cup (I+n))$ ou q_j est la normale intérieure au plan $P_i \quad i \in J$.

Théorème 24 :

Si $q_j^* \quad j \in \bar{J}_0$ est la base de R^n duale de la base $q_j, \quad j \in \bar{J}_0$ alors dans la matrice simplexe $S(J_0)$:

$$\left| \begin{array}{ll} z_{ij} = - \langle q_i, q_j^* \rangle & i \in J_0, \quad j \in \bar{J}_0 \\ d_j = - \langle c, q_j^* \rangle & j \in \bar{J}_0 \\ z_j = - \langle q_j, x(J_0) \rangle + p_j & j \in J_0 \end{array} \right.$$

DEMONSTRATION

a. D'après le théorème 23

$$-q_i = \sum_{j \in \bar{J}_0} z_{ij} q_j \quad i \in J_0$$

par suite

$$z_{ij} = - \langle q_i, q_j^* \rangle$$

de même

$$-c = \sum_{j \in J_0} d_j q_j$$

donc

$$d_j = - \langle c, q_j^* \rangle$$

b. $z_j = \langle q_j, x(\bar{J}_0) \rangle + p_j, \quad j \in J_0$ résulte du théorème 23.

Théorème 25 :

Soit J_0 une base admissible de R^m , $K_1 \cup K_2$ une partition de J_0 telle que :

$$d_j > 0 \quad \forall j \in K_1, \quad d_j \leq 0 \quad \forall j \in K_2$$

si $z_j \geq 0 \quad \forall j \in J_0$ alors, $F_{K_1} = D \cap (\bigcap_{j \in K_1} P_j)$ est une face de D de dimension $|K_2|$ telle que :

$$\langle c, x(\bar{J}_0) \rangle = \inf [\langle c, x \rangle \mid x \in F_{K_1}]$$

$F_{K_2} = D \cap (\bigcap_{j \in K_2} P_j)$ est une face de dimension $|K_1|$ telle que :

$$\langle c, x(\bar{J}o) \rangle = \sup [\langle c, x \rangle, x \in F_{K_2}]$$

DEMONSTRATION :

La variété linéaire $V_{(K_1)} = \bigcap_{j \in K_1} P_j$ est de dimension $n - |K_1| = |K_2|$ car la famille $q_j, j \in K_1$ est libre. Elle ne contient aucun élément intérieur à D puisque tous les plans $P_j, j \in \bar{J}o$ sont des plans d'appui. Il suffit de montrer que V_{K_1} contient $|K_2| + 1$ points indépendants de D pour que $F_1 = V_{K_1} \cap D$ soit une face de dimension $|K_2|$.

Or, il est immédiat que l'on peut choisir $\lambda_i, \forall i \in K_2$ assez petit pour que $x^i = x(\bar{J}o) + \lambda_i q_i^*$ appartienne à D . Les $|K_2| + 1$ points $(x_i \forall i \in K_2, x(\bar{J}o))$ sont indépendants puisque la famille $q_i^* i \in K_2$ est libre. Donc $F_{(K_1)} = V_{(K_1)} \cap D$ est une face de dimension $|K_2|$ de D .

Soit $x \in F_{K_1}$, il est immédiat que x peut être mis sous la forme $x = x(\bar{J}o) + \sum_{j \in K_2} \lambda_j q_j^* \lambda_j \geq 0$.

d'où :

$$\langle c, x \rangle = \langle c, x(\bar{J}o) \rangle + \sum_{j \in K_2} \lambda_j \langle c, q_j^* \rangle = \langle c, x(\bar{J}o) \rangle - \sum_{j \in K_2} \lambda_j d_j$$

d'après la définition de $K_2 \quad d_j \leq 0 \quad \forall j \in K_2$ d'où $\langle c, x \rangle \geq \langle c, x(\bar{J}o) \rangle$

donc $\langle c, x(\bar{J}o) \rangle = \inf [\langle c, x \rangle \mid x \in F_{(K_1)}]$. On voit même que :

$$\langle c, x(\bar{J}o) \rangle = \sup [\langle c, x \rangle \mid x \in F_{(K_2)}]$$

2,24 - CRITIQUE DE LA METHODE SIMPLEXE

On a vu que la méthode simplexe consiste à cheminer de point extrême en point extrême de D à partir d'un point extrême de départ x_0 jusqu'à ce qu'un point extrême solution $\hat{x}$ soit atteint. Ce point x sera caractérisé par la propriété topologique suivante le vecteur $-c$ appartient au cône engendré par les normales intérieures à n plans indépendants contenant x .

Les critiques que l'on peut formuler vis-à-vis de la méthode simplexe sont d'ordre théorique et pratique.

CRITIQUES D'ORDRE THEORIQUE :

1. Mode de cheminement

Soit $B(x, \epsilon)$ une boule ouverte de R^n de centre x et de rayon $\epsilon > 0$, définissons un voisinage d'ordre ϵ de D par $V(\epsilon) = \bigcup_{x \in D} B(x, \epsilon)$. Soit D_1 un polyèdre de R^n tel que :

$$D \subset D_1 \subset V(\epsilon).$$

D_1 sera dit voisin d'ordre ϵ de D .

Pour tout $\epsilon > 0$ arbitrairement petit, on peut choisir D_1 tel que :

- x_0 point extrême de D soit point extrême de D_1 .

- $\hat{x}$ point extrême solution pour D soit point extrême solution pour D_1 .
- le rapport du nombre de points extrêmes de D_1 au nombre de points extrêmes de D soit arbitrairement grand.

On voit donc que pour D_1 presque équivalent à D pour de nombreuses propriétés métriques le rapport du nombre d'étapes dans le cheminement simplicial de x_0 à $\hat{x}$ dans D_1 au nombre d'étapes dans le même cheminement dans D sera arbitrairement grand.

2. Point de départ

Il est nécessaire de connaître x_0 point extrême de D pour initialiser le cheminement simplicial; or, en général, on ne connaît pas de point extrême de D à priori.

On remédie à cet inconvénient en incluant D dans un polyèdre D_1 dont on connaît un point extrême et tel que tout point extrême de D soit point extrême de D_1 .

Cette procédure à l'inconvénient d'augmenter sensiblement le nombre des étapes simpliciales. Dans les cas concrets qu'il nous a été donné de connaître, le nombre d'étapes est fréquemment doublé.

3. Critère de choix en un point extrême

Pour passer d'un point extrême à un autre, il convient de faire rentrer dans la base admissible J_0 indice k à la place d'un indice r . Pour obtenir une augmentation de la fonction économique il faut que $d_k < 0$. Une règle fréquemment utilisée en pratique est de choisir k tel que :

$$d_k = \inf d_i^- \quad (d_i^- \text{ signifiant } d_i < 0).$$

elle ne repose en fait sur aucune considération théorique; en effet, la détermination de l'optimum repose sur le choix d'une base admissible telle que : $d_j \geq 0, \forall j \in J$.

Si J_0 est base admissible pour le problème :

$$\sup [\langle c, x \rangle \mid Ax \leq b, x \geq 0]$$

il est aisé de montrer qu'elle est encore admissible pour le problème :

$$\sup [\langle c', \xi \rangle \mid A' \xi \leq b', \xi \geq 0]$$

avec :

$$A' = D_1 A D_2 \quad b' = D_1 b \quad c' = D_2 c$$

ou D_1 et D_2 sont deux matrices diagonales d'éléments tous positifs. On peut donc modifier arbitrairement la valeur (mais non le signe) des coefficients $d_j, j \in J$ sans influencer sur les propriétés topologiques du polyèdre.

Puisque la valeur des coefficients d_j ne peut donner naissance à un critère de choix, il pourrait paraître intéressant puisque les signes des d_j sont invariants de chercher à chiffrer la « distance » d'un point extrême $x(\bar{J}_0)$ par rapport à un élément optimum en fonction du nombre de coefficients d_j positifs.

En fait, le théorème 25 montre qu'aucune conclusion ne peut être tirée de telles considérations.

Par ailleurs, remarquons que la matrice simplexe ne décrit que des propriétés locales du polyèdre en termes métriques, mais que seules les propriétés topologiques sont utiles.

CRITIQUE D'ORDRE PRATIQUE

La principale critique d'ordre pratique réside dans le fait qu'il faut éviter les pertes de précision au cours des diverses itérations du calcul et de ce fait après un certain nombre d'étapes il convient de recalculer les termes de la matrice simplexe $S(J_0)$ à partir de la matrice simplexe S_0 en effectuant le produit par la matrice du changement de base ce qui nécessite l'inversion d'une matrice $m \times m$.

★

★

★

CHAPITRE III

LA METHODE DU DOUBLE GRADIENT DE M. RAGNAR FRISCH

3,1 - GENERALITES

Il a paru intéressant de consacrer ce chapitre à la méthode du double gradient et aux méthodes qui s'y rattachent pour deux raisons principales :

- a) Cette méthode est la première en date des méthodes de programmation linéaire utilisant un cheminement intérieur au polyèdre défini par les contraintes. Elle est donc résolument différente des méthodes simpliciales.
- b) Certaines réserves avaient été faites par des utilisateurs quant à sa validité.

Nous avons donc tenté de résumer en quelques pages l'ouvrage important de M. Frisch (Principles of linear programming, 219 pages) et de montrer les points délicats relatifs à l'utilisation de la méthode proposée.

Nous utiliserons au maximum les symboles de l'auteur, mais adopterons la notation vectorielle de préférence à la notation développée proposée par l'auteur.

3,2 - SOIT LE PROBLEME :

$$\text{maximer} \quad f(x) = \sum_{i=1}^n x_i p_i$$

$$\begin{aligned} \text{Compte tenu de} \quad & \sum_{i=1}^n a_{ji} x_i \leq b_{j0} & j \in [1 \dots m] \\ & x_i \geq 0 & i \in [1 \dots n] \end{aligned}$$

Les contraintes sont écrites sous la forme :

$$\left| \begin{array}{ll} x_j = b_{j0} + \sum b_{ji} x_i & \text{ou} \quad x_j = b_{j0} + \langle b_j x \rangle \\ x_i \geq 0 & i \in [1 \dots n] \\ x_j \geq 0 & j \in [1 \dots m] \end{array} \right.$$

Elles définissent un polyèdre C d'intérieur $\overset{\circ}{C}$

3,21 - DEPLACEMENT A PARTIR D'UN POINT INTERIEUR A C

Soit $x_0 \in \overset{\circ}{C}$, M. Frisch propose d'effectuer à partir de x_0 le déplacement λd avec :

$$(1) \quad d = p + \mu V$$

$$\text{et} \quad \left| \begin{array}{l} p = (p_1 \dots p_n) = \text{grad } f(x) \\ V = (V_1 \dots V_m) = \text{grad } U \\ U = \sum_{i=1}^n \log x_i + \sum_{j=1}^m \log x_j \end{array} \right.$$

VARIATION DES COMPOSANTES x_i ET x_j DANS CE DEPLACEMENT

de $x = x_0 + \lambda d$, on tire :

$$x_i = x_{i0} + \lambda d_i \quad i \in (1 \dots n),$$

et de :

$$x_j = b_{j0} + \langle b_j, x_0 + \lambda d \rangle = b_{j0} + \langle b_j, x_0 \rangle + \lambda \langle b_j, d \rangle \quad j \in (1 \dots m)$$

en posant :

$$x_{j0} = b_{j0} + \langle b_j, x_0 \rangle \quad \text{et} \quad d_j = \langle b_j, d \rangle \quad \text{on obtient :}$$

$$x_j = x_{j0} + \lambda d_j$$

soit pour $i \in (1, \dots, m+n)$:

$$(2) \quad x_i = x_{i0} + \lambda d_i$$

Par ailleurs, de (1) on tire :

$$d_i = p_i + \mu V_i \quad i \in (1 \dots n)$$

et pour $j \in (1 \dots m)$ puisque l'on a posé $d_j = \langle b_j, d \rangle$

$$d_j = \langle b_j, p \rangle + \mu \langle b_j, V \rangle$$

soit en posant :

$$p_j = \langle b_j, p \rangle$$

$$V_j = \langle b_j, V \rangle$$

$$(3) \quad d_i = p_i + \mu V_i \quad i \in (1, \dots, m+n)$$

VARIATION DE LA FONCTION ECONOMIQUE

$$f_0 = \langle p, x_0 \rangle$$

$$f = \langle p, x \rangle$$

$$\text{d'où : } f - f_0 = \lambda \langle p, d \rangle = \lambda [\langle p, p \rangle + \mu \langle p, V \rangle]$$

ou

$$(4) \quad \Delta f = \lambda (P + \mu M)$$

CHOIX DE λ ET μ

Si $P + \mu M > 0$, Δf sera positif si $\lambda > 0$.

Pour satisfaire à $x_i \geq 0$, $i \in (1 \dots m+n)$, il faudra que, d'après (2) :

$$\left| \begin{array}{l} \lambda \leq \inf_{i \in I^-} \left(- \frac{x_{i0}}{d_i} \right) \\ I^- = \{ i ; d_i < 0, i \in (1, \dots, m+n) \} \end{array} \right.$$

Si $P + \mu M < 0$, Δf sera positif si $\lambda < 0$.

Pour satisfaire à $x_i \geq 0$, $i \in (1, \dots, m+n)$, il faudra que :

$$\lambda \leq - \inf_{i \in I} + \left(\frac{x_{i_0}}{d_i} \right)$$

$$I^+ = [i; d_i > 0, i \in (1, \dots, m+n)]$$

Pour discuter le choix de λ et μ , M. R. Frisch écrit 4 sous la forme :

$$(5) \quad \frac{1}{\Delta f} = \frac{\frac{1}{\lambda}}{P + \mu M} \quad \text{avec } \lambda \neq 0$$

Le maximum de Δf est donc une fonction de μ qui s'écrit :

$$\frac{1}{\Delta f} = \frac{1}{|P + \mu M|} \sup_{i \in I} \left| \frac{d_i}{x_{i_0}} \right|$$

$$\begin{aligned} I^- &= I^+ & \text{si} & P + \mu M > 0 \\ I^- &= I^+ & \text{si} & P + \mu M < 0 \end{aligned}$$

D'après la relation 3 : $d_j = p_j + \mu V_j$ $j \in (1, \dots, m+n)$

on a :

$$(6) \quad \frac{1}{\Delta f} = \frac{1}{|P + \mu M|} \sup_{i \in I \sim} \left| \frac{p_i}{x_{i0}} + \mu \frac{v_i}{x_{i0}} \right|$$

ETUDE GRAPHIQUE DE $\frac{1}{\Lambda_f}$ EN FONCTION DE μ

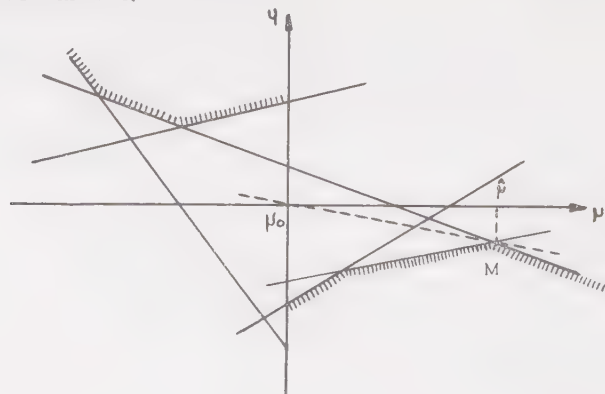
M. R. Frisch trace dans le plan y, μ les droites $y = \frac{p_i}{x_{i0}} + \mu \frac{V_i}{x_{i0}}$

Si M est positif, le rapport $\mu_0 = - \frac{P}{M}$ sera négatif et on aura :

$$P + \mu M < 0 \quad \text{pour} \quad \mu \in [-\infty, \mu_0[$$

$$P + M > 0 \quad \text{pour} \quad \mu \in]\mu_0, +\infty]$$

- Dans l'intervalle $[-\infty, \mu_0[$, il faut considérer l'enveloppe supérieure des demi-droites d'ordonnées positives définies pour $\mu < \mu_0$.
- Dans l'intervalle $]\mu_0, \infty]$, il faut considérer l'enveloppe inférieure des demi-droites d'ordonnées négatives définies pour $\mu > \mu_0$.
- D'après l'expression (6), la valeur $\hat{\mu}$ assurant la plus grande augmentation de f sera celle correspondant au plus petit angle formé par l'axe des μ et le segment $\mu_0 M$, M parcourant les enveloppes définies précédemment.



Si $x_0 \in \bar{C}$, il y a toujours une valeur $\hat{\mu}$ donnant une augmentation optimum à f (il peut y en avoir 2 au plus). $\hat{\mu}$ peut être fini ou infini.

Si $\hat{\mu}$ est fini, deux variables x_i , $i \in (1, \dots, m+n)$ au moins s'annulent pour cette valeur de $\hat{\mu}$.

Le cas $\hat{\mu}$ infini doit être écarté de cette analyse graphique puisque, dans ce cas, $\lambda = 0$ et l'écriture de (4) sous la forme (5) n'est plus admissible.

3,22 - DEPLACEMENTS ULTERIEURS

Le point atteint par le déplacement précédent a conduit à l'annulation de 2 variables

x_i , $i \in (1, \dots, m+n)$ au moins, donc à un point sur la frontière de C .

Ce point est considéré comme un nouveau point x_0 et M. R. Fisch propose la méthode suivante comportant plusieurs cas :

3,221 - Le nouveau déplacement $\lambda(p + \mu V)$ est défini en ne faisant pas entrer dans le calcul de V les variables x_{i_0} nulles $i_0 \in (1, \dots, m+n)$.

Soit $I_0 = [i \mid x_{i_0} = 0]$

Le déplacement $\lambda(p_i + \mu V_i)$ doit être positif pour tout $i \in I_0$, $\lambda > 0$ conviendra si :

$$\sup_{i \in I_1} \left(\frac{-p_i}{V_i} \right) \leq \mu \leq \inf_{i \in I_2} \left(\frac{-p_i}{V_i} \right)$$

$$I_1 = [i \in I_0; V_i > 0]$$

$$I_2 = [i \in I_0; V_i < 0]$$

$\lambda < 0$ conviendra si :

$$\sup_{i \in I_2} \left(\frac{-p_i}{V_i} \right) \leq \mu \leq \inf_{i \in I_1} \left(\frac{-p_i}{V_i} \right)$$

Il suffira de placer sur l'axe de μ les 4 bornes précédentes pour voir si f peut être augmentée à l'aide de la méthode graphique présentée ci-dessus.

Il peut se faire évidemment qu'aucun intervalle convenable pour μ ne puisse être mis en évidence. Dans ce cas, M. R. Frisch propose 3 méthodes dont 2 sont incorrectes, comme nous le montrerons plus loin, et la 3ème consiste à résoudre un nouveau programme linéaire.

3,222 - PREMIERE METHODE

Deux variables seulement sont nulles : x_1 et x_2 par exemple.

Le déplacement sera λb , λ positif de telle sorte que b soit dirigé vers l'intérieur de C et contenu dans le plan des normales b_1 et b_2 aux plans $x_1 = 0$, $x_2 = 0$, normales orientées vers l'intérieur de C .

Ce déplacement doit assurer à x_1 et x_2 de ne pas devenir négatives.

Soit $b = u_1 b_1 + u_2 b_2$ avec $u_1^2 + u_2^2 = 1$

Il est nécessaire et suffisant pour que x_1 et x_2 soient positives ou nulles que :

$$\begin{aligned} \langle b, b_1 \rangle &\geq 0 \\ \langle b, b_2 \rangle &\geq 0 \end{aligned}$$

Soit :

$$(7) \quad \begin{cases} u_1 < b_1, b_2 > + u_2 < b_2, b_2 > \geq 0 \\ u_1 < b_1, b_2 > + u_2 < b_2, b_2 > \geq 0 \end{cases}$$

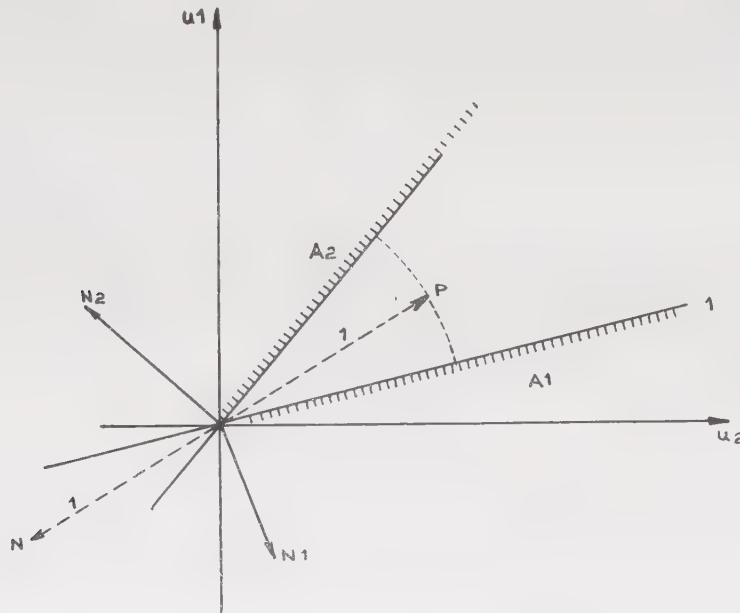
Comme $< b_1, b_1 >$ et $< b_2, b_2 >$ sont positifs et $< b_1, b_2 >^2 - < b_1, b_1 > < b_2, b_2 > \leq 0$ (inégalité de Schwarz), il existe toujours des solutions au système (7).

La variation de f sera alors :

$$\Delta f = < \lambda b, p > = \lambda [u_1 < b_1, p > + u_2 < b_2, p >]$$

Elle doit être rendue maximum compte tenu de (7) et de $u_1^2 + u_2^2 = 1$.

M. R. Fisch considère les normales N_1 et N_2 aux droites (1) et (2) (voir figure)



et, sur ces droites, les points A_1 et A_2 tels que $u_1^2 + u_2^2 = 1$ puis il prend le vecteur P ($< b_1, p >$, $< b_2, p >$) et dit, page 116 dixième ligne avant la fin : « If the vector P points along the line (2), the optimum point will be A_2 . As P turns counterclockwise from his position, the result is as follows : if P points anywhere in the sector between A_2 and N_2 , the optimum point will still be in A_2 and this applied even until the direction of P has reached that of N (N situé sur la médiatrice de $A_1 A_2$ à une distance 1 de 0). As the direction of P passes that of N , the optimum point jumps from A_2 to A_1 and rests there until the direction of P points along the line (1). If the direction of P is between that of the line (1) and (2) the optimum point is simply the intersection between P and the part of the unit circle ».

Cette procédure est manifestement erronée car, supposons que :

1. Les deux droites 1 et 2 forment un angle aigu (il suffira que $< b_1, b_2 >$ soit négatif).
2. $< p, b_1 >$ et $< p, b_2 >$ sont négatifs.

Dans ce cas, $\Delta f = \lambda [u_1 < p, b_1 > + u_2 < p, b_2 >]$ ne peut être positive puisque λ, u_1, u_2 sont astreintes à être positives.

Il est d'ailleurs aisé de voir que l'augmentation la plus grande de f aurait lieu dans la variété intersection des plans $x_1 = 0$ et $x_2 = 0$.

3,223 - DEUXIEME METHODE

Les variables $x_i, i \in I_0$ sont nulles.

Dans le cas où $I_0 = [1, 2]$, on a un déplacement b qui satisfait à :

$$\langle b, b_1 \rangle \geq 0$$

$$\langle b, b_2 \rangle \geq 0$$

$$\text{en prenant } b = \frac{1}{\sqrt{\langle b_1, b_1 \rangle}} \cdot b_1 + \frac{1}{\sqrt{\langle b_2, b_2 \rangle}} \cdot b_2$$

c'est une conséquence de l'inégalité de Schwarz.

M. R. Frisch propose de généraliser cette méthode permettant d'obtenir un déplacement vers l'intérieur de C . Il dit, page 117 (huitième ligne) :

« We want to construct a vector b in x space, that as a practical rule in work with the double gradient method and be taken as pointing from x_0 as much as possible «into the middle» of the sector, where all the variables $x_i (i \in I_0)$ are positive ».

Et, plus loin, il ajoute (page 118, deuxième ligne) : « Heuristically, we may conjecture that this (la relation valable pour $I_0 = [1, 2]$) is true in the case of v arbitrary » (v = nombre de variables x_i nulles).

L'auteur propose de considérer un déplacement :

$$b = \sum_{i \in I_0} u_i b_i \quad \text{avec} \quad u_i = \frac{1}{\sqrt{\langle b_i, b_i \rangle}}$$

Nous remarquons qu'il faut que le déplacement ait lieu dans le bon sens à partir des plans :

$$x_i = 0, \quad i \in I_0$$

Ce qui conduit à écrire :

$$\sum_{j \in I_0} \frac{\langle b_j, b_j \rangle}{\sqrt{\langle b_j, b_j \rangle}} \geq 0, \quad \forall i \in I_0$$

Il est aisé de construire un contre exemple où ces relations ne sont pas vérifiées.

Considérons les 3 contraintes :

$$\begin{array}{l|l} (1) & x_3 - \frac{1}{2} \geq 0 \\ (2) & \sqrt{3} x_1 - x_3 + \frac{1}{2} - \sqrt{3} \geq 0 \\ (3) & 4x_1 + 4x_2 - 3\sqrt{2} x_3 - 8 + 3\sqrt{2} \geq 0 \end{array}$$

Ces contraintes ne sont pas incompatibles ; par exemple $x = (3, 2, 2)$ est un élément acceptable.

En introduisant les variables d'écart n_1, n_2, n_3 ces 3 contraintes s'écrivent :

$$n_1 = x_3 - \frac{1}{2}$$

$$n_2 = \sqrt{3} x_1 - x_3 + \frac{1}{2} - \sqrt{3}$$

$$n_3 = 4x_1 + 4x_2 - 3\sqrt{2} x_3 - 8 + \frac{3\sqrt{2}}{2}$$

Supposons être arrivés au point $x_1 = x_2 = 1$ $x_3 = \frac{1}{2}$

Dans ce cas : $n_1 = 0$, $n_2 = 0$, $n_3 = 0$

Le déplacement recommandé par M. R. Frisch est :

$$b = \frac{b_1}{\sqrt{\langle b_1, b_1 \rangle}} + \frac{b_2}{\sqrt{\langle b_2, b_2 \rangle}} + \frac{b_3}{\sqrt{\langle b_3, b_3 \rangle}}$$

Il doit satisfaire en particulier à $\langle b, b_1 \rangle \geq 0$, soit à :

$$\sqrt{\langle b_1, b_1 \rangle} + \frac{\langle b_2, b_1 \rangle}{\sqrt{\langle b_2, b_2 \rangle}} + \frac{\langle b_3, b_1 \rangle}{\sqrt{\langle b_3, b_3 \rangle}} \geq 0$$

Or, $b_1 = (0, 0, 1)$, $b_2 = (\sqrt{3}, 0, -1)$, $b_3 = (4, 4, -3\sqrt{2})$

ce qui donne :

$$1 - \frac{1}{\sqrt{1+3}} - \frac{3\sqrt{2}}{\sqrt{16+16+18}} = 1 - \frac{1}{2} - \frac{3}{5} = -\frac{1}{10}$$

valeur non admissible.

M. R. Frisch poursuit d'ailleurs en disant (page 118, ligne 9) :

« Whatever the definition of b , we may put :

$$d = p + \mu b$$

and make an optimum determination of μ by standard method of figure (12.31) (c'est-à-dire par la méthode graphique proposée) used with the appropriate side of condition for μ ».

Or, soit A l'espace vectoriel affine contenant x_0 [$x_i = 0$, $i \in I_0$, $x_i \neq 0$, $i \in I_0$] et engendré par les vecteurs p et b .

Il est aisé de voir que la méthode proposée ne sera pas valable dans les 2 cas suivants :

$$(1) \quad A \cap C = [x_0], \quad \sup_{x \in C} f(x) \neq f(x_0)$$

$$(2) \quad \sup_{x \in (A \cap C)} f(x) = f[x_0] \neq \sup_{x \in C} f(x)$$

Nous avons construit un exemple où le premier cas :

$$A \cap C = [x_0], \quad \sup_{x \in C} f(x) \neq f(x_0), \quad \text{se produit:}$$

Soit à minimiser : $-x_1 + 10 x_2$
compte tenu des contraintes

$$x_3 - \frac{1}{2} \geq 0$$

$$\sqrt{3} x_1 - x_3 + \frac{1}{2} - \sqrt{3} \geq 0$$

$$4x_1 + 4x_2 - 3\sqrt{2}x_3 - 8 + \frac{3\sqrt{2}}{2} \geq 0$$

Le point $x_0 (1, 1, 1/2)$ vérifie strictement les 3 premières équations, donc annule les 3 variables d'écart n_1, n_2, n_3 .

Puisque $n_1 = n_2 = n_3 = 0$, on a :

$$U = \log x_1 + \log x_2 + \log x_3$$

d'où :

$$V_{x_1} = 1$$

$$V_{x_2} = 1$$

$$V_{x_3} = 2$$

et :

$$V_{n_1} = \langle b_1, V \rangle = 2$$

$$V_{n_2} = \langle b_2, V \rangle = \sqrt{3} - 2$$

$$V_{n_3} = \langle b_3, V \rangle = 8 - 6\sqrt{2}$$

puis :

$$p_{n_1} = \langle b_1, p \rangle = 0$$

$$p_{n_2} = \langle b_2, p \rangle = -\sqrt{3}$$

$$p_{n_3} = \langle b_3, p \rangle = 36$$

Appliquons la méthode décrite en 3,221.

Un déplacement $\lambda (p + \mu V)$ sera possible pour $\lambda \geq 0$

$$\text{si } 0 \leq \mu \leq \frac{-\sqrt{3}}{2 - \sqrt{3}}$$

donc un déplacement avec λ positif n'est pas possible.

Un déplacement avec λ négatif sera possible si :

$$\frac{36}{6\sqrt{2} - 8} \leq \mu \leq 0$$

condition également impossible à satisfaire.

Donc, la méthode proposée en 3,221 ne convient pas. Que peut-on dire de la méthode proposée en 3,222 ?

Au point $x = (1, 1, 1/2)$, les arêtes d'intersection des plans $P_1 - P_2 - P_3$ deux à deux correspondant aux contraintes 1, 2, 3, sont dirigées suivant V_1, V_2, V_3 , tels que

$$\begin{array}{ll} V_1 = (0, 1, 0) & \text{parallèle à } P_1 \cap P_2 \\ V_2 = (1, -1, 0) & \text{parallèle à } P_1 \cap P_3 \\ V_3 = (4, 3\sqrt{6} - 4, 4\sqrt{3}) & P_2 \cap P_3 \end{array}$$

Le vecteur $b = \frac{b_1}{\sqrt{\langle b_1, b_1 \rangle}} + \frac{b_2}{\sqrt{\langle b_2, b_2 \rangle}} + \frac{b_3}{\sqrt{\langle b_3, b_3 \rangle}}$

est dirigé suivant $U = (5\sqrt{3} + 4\sqrt{2}, 4\sqrt{2}, -1)$.

Le déplacement $\lambda (p + \mu b)$ aura lieu dans le plan contenant l'élément $(1, 1, 1/2)$ et de normale proportionnelle à $p \wedge b$, de composantes $(-10, -1, -44\sqrt{2} - 50\sqrt{3})$.

Les produits scalaires entre $p \wedge b$ et V_1, V_2, V_3 sont :

$$\langle v_1, p \wedge b \rangle = -1$$

$$\langle v_2, p \wedge b \rangle = -9$$

$$\langle v_3, p \wedge b \rangle = -636 - 179\sqrt{6}$$

Tous ces produits sont négatifs ; par suite, le plan dans lequel s'effectuera le déplacement n'a en commun avec C que le point $(1, 1, 1/2)$ et, de ce fait, ne convient pas.

La fonction économique n'est d'ailleurs pas minimum en ce point puisqu'un déplacement suivant V_2 entraîne une variation $\Delta f = \langle p, V_2 \rangle = -9$.

3,224 - TROISIEME METHODE

Au lieu d'écrire $b = \sum_{i \in I_0} \frac{b_i}{\sqrt{\langle b_i, b_i \rangle}}$ M. R. Frisch propose de considérer (voir page 118)

$$b = \sum_{i \in I_0} u_i \cdot b_i$$

La méthode précédente est donc un cas particulier de celle-ci. On aura alors un déplacement :

$$\lambda d = \lambda (p + \mu b) = \lambda [p + \sum_{i \in I_0} V_i b_i] \quad \text{avec} \quad V_i = \mu u_i$$

Le nouveau point x sera :

$$x = x_0 + \lambda d = x_0 + \lambda [p + \sum_{i \in I_0} V_i b_i]$$

soit pour $j \in (1, \dots, m+n)$

$$x_j = x_{j0} + \lambda [\langle b_j, p \rangle + \sum_{i \in I_0} \langle b_j, b_i \rangle V_i]$$

Pour $\forall j \in I_0$, on devra avoir en particulier :

$$\lambda [\langle b_j, p \rangle + \sum_{i \in I_0} \langle b_j, b_i \rangle V_i] \geq 0$$

La variation de la fonction économique sera :

$$\Delta f = \lambda [\langle p, p \rangle + \sum_{i \in I_0} \langle p, b_i \rangle V_i]$$

En prenant $\lambda > 0$, on aura à résoudre le programme linéaire suivant :

$$\text{maximer } \sum_{i \in I_0} \langle p, b_i \rangle V_i$$

$$\text{compte tenu de } \sum_{i \in I_0} \langle b_i, b_j \rangle V_i + \langle b_j, p \rangle \geq 0 \quad \forall j \in I_0$$

Nous démontrons ici que, contrairement aux autres méthodes proposées par M. R. Frisch, cette procédure permettra d'augmenter f .

En effet, soient dans la base canonique de R^n :

P la matrice colonne représentant p

B la matrice représentant les vecteurs $b_i \quad i \in I_0$ de rang $|L_0|$

V la matrice colonne composée des $V_i \quad i \in I_0$

$$\text{Les contraintes } \sum_{i \in I_0} \langle b_j, b_i \rangle V_i + \langle b_j, p \rangle \geq 0 \quad \forall j \in I_0$$

s'expriment sous la forme :

$$({}^t B B) V + {}^t B P = h \quad h = \text{matrice colonne de termes } h_i \geq 0$$

$$\text{d'où } V = ({}^t B B)^{-1} (h - {}^t B P)$$

$$\text{et } \Delta f = \langle p, p \rangle + \sum_{i \in I_0} \langle p, b_i \rangle V_i, \text{ s'écrit sous la forme :}$$

$$\Delta f = {}^t P P + {}^t P B V$$

$$\text{soit } \Delta f = {}^t P [I - B ({}^t B B)^{-1} {}^t B] P + {}^t P B ({}^t B B)^{-1} h$$

La matrice $M = I - B ({}^t B B)^{-1} {}^t B$ est idempotente.

Par suite, ses valeurs propres sont 0 ou 1,

$$\text{ce qui fait que } {}^t P M P \geq 0$$

et, par suite $\Delta f \geq 0$.

Le problème se pose alors de savoir comment résoudre le programme linéaire. En opérant par la méthode du double gradient, on risque évidemment de se retrouver dans une situation analogue et de devoir à résoudre un nouveau programme linéaire. Cette situation semble avoir échappée à M. R. Frisch qui écrit page 119 :

« To get a simple device for application in a step by step approximation, we would have to stop short of this complete problem and rely, for instance, on some such simple form as

$$b = \sum_{i \in I_0} \frac{b_i}{\sqrt{\langle b_i, b_i \rangle}}, \text{ taken in conjunction with } d = p + \mu b \text{ where } \mu \text{ is a sin-}$$

gle parameter to be disposed of ».

Par conséquent, M. R. Frisch renonce à utiliser cette troisième méthode qui seule donne l'assurance de sortir des « pièges » provoqués par les variables nulles.

3,3 - ESSAI DE RESOLUTION D'UN PROGRAMME LINEAIRE PAR INTEGRATION D'EQUATION DIFFERENTIELLES DERIVANT DE LA METHODE DE M. R. FRISCH

Il peut être intéressant de regarder s'il serait possible de trouver la solution d'un programme linéaire par intégration d'un système d'équations différentielles dérivant de la méthode du double gradient.

Considérons le point $x(x_1, \dots, x_n)$ comme se déplaçant sous l'effet de la force : $p + \lambda \text{ grad } U$.

On aura donc :

$$\dot{x} = p + \lambda \text{ grad } U$$

$$\dot{x} = \frac{dx}{dt}$$

$$U = \sum_1^n \log x_i + \sum_1^m \log x_j$$

Prenons l'exemple élémentaire suivant :

maximiser $f(x) = x_1 + x_2$

sous les contraintes

$$\left| \begin{array}{l} x_1 \leq \alpha \\ x_2 \leq \alpha \\ x_1 \text{ et } x_2 \geq 0 \end{array} \right.$$

Il est évident que la solution s'obtient pour $x_1 = x_2 = \alpha$.

Par la méthode différentielle, on aura à intégrer :

$$\frac{dx_1}{dt} = 1 + \lambda \frac{\partial U}{\partial x_1}$$

$$\frac{dx_2}{dt} = 1 + \lambda \frac{\partial U}{\partial x_2}$$

$$U = \log x_1 + \log x_2 + \log (\alpha - x_1) + \log (\alpha - x_2)$$

soit :

$$\frac{dx_1}{dt} = 1 + \lambda \left(\frac{1}{x_1} - \frac{1}{\alpha - x_1} \right)$$

$$\frac{dx_2}{dt} = 1 + \lambda \left(\frac{1}{x_2} - \frac{1}{\alpha - x_2} \right)$$

L'équation différentielle

$$\frac{dx}{dt} = 1 + \lambda \left(\frac{1}{x} - \frac{1}{\alpha - x} \right)$$

a pour solution :

$$x + \lambda I = t + C \quad (C = \text{constante définie par les conditions initiales})$$

$$I = \int \frac{\alpha - 2x}{x^2 + x(2\lambda - \alpha) - \lambda\alpha} dx$$

L'étude de l'hyperbole $H(x\lambda) \equiv x^2 + x(2\lambda - \alpha) - \lambda\alpha = 0$, pour α donné > 0 montre que pour $\lambda \geq 0$ la racine x_1 sera comprise entre α et $\alpha/2$, x_2 étant négative ; pour $\lambda \leq 0$ la racine x_2 sera comprise entre 0 et $\alpha/2$, x_1 étant positive.

Ces racines correspondent d'ailleurs à des positions où la force $f = 1 + \left(\frac{1}{x} - \frac{1}{\alpha - x} \right) \lambda$ s'exerçant sur le point mobile s'annule.

Considérons l'équilibre en ces positions :

$$f = 1 + \lambda \left(\frac{1}{x} - \frac{1}{a-x} \right) = - \frac{x^2 + x(2\lambda - a) - \lambda a}{x(a-x)} = - \frac{(x-x_1)(x-x_2)}{x(a-x)}$$

Si $\lambda \geq 0$: on a $x_1 - x_2 \geq 0$ d'où $f \geq 0$ pour $x \leq x_1$ et $f \leq 0$ pour $x \geq x_1$. La position x_1 est donc une position d'équilibre stable.

Si $\lambda \leq 0$: on a $x_1 - x_2 \geq 0$ et $x_1, x_2 \geq 0$ d'où $f \geq 0$ pour $x \leq x_2$. La position x_2 est donc une position d'équilibre instable. D'ailleurs, ce cas n'a pas d'intérêt car le point mobile est attiré par les faces du polyèdre au lieu d'être repoussé et n'est donc plus astreint à rester dans le polyèdre.

ETUDE DU MOUVEMENT

(1) $\lambda > 0$ $x^* < x_1$: on a, en mettant en évidence tous les termes positifs :

$$x - \lambda \left(\frac{2x_1 - a}{x_1 - x_2} L(x_1 - x) + \frac{a - 2x_2}{x_1 - x_2} L(x - x_2) \right) = t + C$$

La position x_1 sera atteinte au bout d'un temps infini.

(2) $\lambda > 0$ $x > x_1$: on a :

$$x - \lambda \left(\frac{2x_1 - a}{x_1 - x_2} L(x - x_1) + \frac{a - 2x_2}{x_1 - x_2} L(x - x_2) \right) = t + C$$

Même conclusion que précédemment.

Par conséquent, l'intégration des équations différentielles de départ ne peut conduire à l'optimum, sauf ici, dans le cas trivial $\lambda = 0$, et C (constante d'intégration) = 0

Cette situation visiblement se produirait aussi dans le cas général.

Au surplus, l'intégration numérique s'avèrerait extrêmement délicate pour les éléments x voisins de la frontière du polyèdre, donc en particulier au voisinage de l'optimum car, dans ce cas, $\| \text{grad } u \|$ est très élevée. Il est évident qu'il serait impossible de maintenir λ constant.

3.4 - METHODE DE M. G. PARISOT

Cette idée de variation de λ a été exploitée et généralisée par M. G. Parisot (comptes rendus de de l'Académie des Sciences, tome 248, pages 1.470, 1.472 ; séance du 9 mars 1959).

M. Parisot propose d'intégrer l'équation vectorielle :

$$\begin{cases} \dot{x} = \lambda [p + \text{grad } U] \\ U = \sum_{i=1}^n a_i \log x_i + \sum_{j=1}^{n+m} a_j \log x_j \\ x_j = b_{j0} + \langle b_j, x \rangle \end{cases}$$

Il écrit (C.R.A.S. déjà cité) :

« le vecteur (a_i) , objectif de précision, (sera) ajusté en cours de calcul pour n'être ni trop,

ni trop peu ambitieux, et un pas d'intégration λ (sera) déterminé par les règles classiques en calcul intégral. La méthode du Professeur Ragnar Frisch peut être considérée comme une application particulière de cette formule $\dot{x} = \lambda [p + \text{grad } u]$.

- en prenant tous les α_i égaux entre eux (leur valeur commune étant α),
- en optimisant d'une certaine façon, à chaque itération, les deux coefficients α et λ .

Ces règles semblent passablement arbitraires et restrictives, elles simplifient apparemment les calculs, mais diminuent la convergence ».

3,5 - AUTRE METHODE SE RATTACHANT A LA METHODE DU DOUBLE GRADIENT

M. Peuchot envisage la méthode suivante (présentée au Congrès International de Recherche Opérationnelle, Aix-en-Provence 1960).

Un point M_0 intérieur au convexe C ayant été trouvé, M. Peuchot fait passer l'hyperplan représentatif de la fonction économique par le point M_0 , puis décolle légèrement celui-ci dans le sens

$p = \text{grad } f(x)$. Le point M_0 se déplace sous l'action du vecteur $\text{grad } u$, $u = \sum_{i=0}^{m+n} \log x_i$ (la sommation est étendue à la variable d'écart relative à la fonction économique x_0 , à toutes les variables x_i , $i \in (1, \dots, n)$ définissant l'élément M , et à toutes les variables d'écart x_j , $j \in (n+1, \dots, n+m)$).

M. Peuchot démontre que le point M_0 s'arrêtera en un point Ω_1 de potentiel maximum.

Ce point Ω_1 sera considéré comme un nouveau point M_0 et on appliquera de nouveau le processus.

Il semble en effet que l'on puisse obtenir de la sorte une suite infinie de points Ω_i convergeant vers l'optimum.

En conclusion de cet exposé critique de la méthode de M. R. Frisch avec mention des méthodes qui s'y rattachent, nous avons montré que la méthode du double gradient, malgré l'intérêt qu'elle semblait présenter de prime abord, était difficilement applicable par suite de ses graves lacunes quand la frontière du polyèdre est atteinte.

Les deux méthodes qui s'y rattachent, par contre, permettent probablement d'aboutir à l'optimum.

La première exposée, celle de M. G. Parisot, nécessite un choix judicieux mais sans doute délicat du vecteur α .

La deuxième, celle de M. Peuchot, présente l'inconvénient de calculs assez lourds pour obtenir les points Ω_i et, de plus, rien n'est dit sur la rapidité de convergence de la suite des Ω_i .

★

★

★

METHODE SEQUENTIELLE INTERIEURE

4,1 - GENERALITES ET PRINCIPE DE LA METHODE

- Soit le problème de programmation linéaire

$$\sup [\langle c, x \rangle \mid x \in D \subset \mathbb{R}^n]$$

$$D = \bigcap_{j \in J} H_j \quad J = [1, \dots, m+n]$$

H_j demi-espace fermé de $\mathbb{R}^n$ défini par

$$[x \mid \langle a_j, x \rangle + b_j \geq 0]$$

Si pour $j \in N = [1, \dots, n]$ on a $H_j = [x \mid x_j \geq 0]$ le problème posé est équivalent au programme linéaire canonique : $\sup [\langle c, x \rangle \mid Ax \leq b, x \geq 0]$

avec $A =$ matrice $m \times n$ constituée par les lignes $-a_j \quad j \in J - N$

$b =$ matrice $m, 1$ d'élément $b_j \quad j \in J - N$

- Supposons que :

1. L'intérieur $\overset{\circ}{D}$ de D soit non vide et que $x_0 \in \overset{\circ}{D}$

2. On puisse construire une suite croissante de variétés linéaires $[x_0] = V_0 \subset V_1 \subset V_2 \dots \subset V_k \dots$ de dimension respectivement $0, 1, 2, \dots, k ; k \leq n$, et que dans chaque variété on puisse trouver x_i tel que :

$$\langle c, x_i \rangle = \sup [\langle c, x \rangle \mid x \in V_i \cap D]$$

Alors, comme $V_n \cap D = D$, x_n sera l'optimum atteint en au plus n étapes.

Définition :

On dira que la famille de vecteurs de $\mathbb{R}^n$, $[x_k] \quad k \in K$, est en position générale dans $\mathbb{R}^n$ si toute sous famille $x_p, \quad p \in K_0 \subset K, \quad |K_0| \leq n$ est libre dans $\mathbb{R}^n$.

HYPOTHESE 1 :

On supposera provisoirement que :

la famille $[c \cup a_j : \forall j \in J]$ est en position générale dans $\mathbb{R}^n$.

En conséquence, le vecteur c ne peut être combinaison linéaire de moins de n vecteurs $a_j, j \in J$.

Théorème 26 :

x_p élément de la suite appartient à au moins p plans $P_j, j \in J$ dont les normales a_j forment une famille de vecteurs linéairement indépendants.

DEMONSTRATION :

a) $D_p = V_p \cap D$ est un polyèdre de dimension p .

En effet, $D_p = V_p \cap (\bigcap_{j \in J} H_j) = \bigcap_{j \in J} (V_p \cap H_j)$

$V_p \cap H_j$ est, soit un demi-espace fermé h_j de V_p limité par le plan $\Pi_j = V_p \cap P_j$, soit V_p tout entier.

Donc, $D_p = \bigcap_{j \in J} h_j$ est un tronçon de V_p . Comme D est borné, D_p l'est aussi, donc D_p est un polyèdre, et puisque V_p est de dimension p , D_p est un polyèdre de dimension p .

b) x_p rendant maximum $\langle c, x \rangle$ dans D_p est un élément extrême de D_p , donc est déterminé par l'intersection de p plans Π_j , $j \in J_0 \subset J$, dont les normales forment une base du sous-espace vectoriel E_p parallèle à V_p tel que la projection orthogonale $-\bar{c}$ de $-c$ sur E_p admet une décomposition positive. Les normales à Π_j , $j \in J_0$ étant les projections orthogonales des normales à P_j , $j \in J_0$ sur E_p , les normales α_j , $j \in J_0$ forment une famille libre dans R^n .

Comme $\Pi_j \subset P_j$, $\forall j \in J$, x_p appartient à au moins p plans P_j dont les normales α_j forment une famille de vecteurs linéairement indépendants.

x_p peut, dans certains cas, appartenir à plus de p plans P_j . Nous ferons provisoirement l'hypothèse 2.

HYPOTHESE 2 :

x_p appartient à exactement p plans P_j .

4.2 - CONSTRUCTION DE LA SUITE $x_0, x_1, \dots, x_n$

a) x_0 appartient à l'intérieur $\overset{\circ}{D}$ de D . Il est naturel de définir $V_1 = [x | x = x_0 + \lambda_0 c]$ c'est-à-dire une variété parallèle au gradient de la fonction à maximiser. Le polyèdre $D_1 = V_1 \cap D$ et de dimension 1, donc x_1 appartient d'après l'hypothèse 2 à un seul plan P_j , $j \in J$. On peut, sans restreindre la généralité, supposer que $j = 1$.

b) On définira $V_2 = [x | x = x_0 + \lambda_0 c + \lambda_1 \alpha_1]$ qui, d'après l'hypothèse 1 est de dimension 2. V_2 contient V_1 et l'élément x_2 maximant $\langle c, x \rangle$ dans $D_2 = V_2 \cap D$ appartient, d'après l'hypothèse 2, à 2 plans seulement P_{j_1} et P_{j_2} . L'un au moins des 2 indices j_1, j_2 est différent de 1; on peut, sans restreindre la généralité, poser $j_1 = 2$.

c) Plus généralement, si x_p , $p \leq n - 1$ maximise $\langle c, x \rangle$ dans $D_p = V_p \cap D$ avec $V_p = [x | x = x_0 + \lambda_0 c + \dots + \lambda_p \alpha_p]$ de dimension p ; d'après l'hypothèse 1, x_p appartient d'après l'hypothèse 2 à p plans seulement $P_{j(1)} \dots P_{j(p)}$ l'un au moins des indices $j(1) \dots j(p)$ est différent de 1, ..., $p - 1$ et on peut poser $j(p) = p$ sans restreindre la généralité; on définira alors V_{p+1} :

$$V_{p+1} = [x | x = x_0 + \lambda_0 c + \lambda_1 \alpha_1 + \dots + \lambda_p \alpha_p]$$

V_{p+1} contient V_p et, d'après l'hypothèse 1, est de dimension $p + 1$.

En continuant de cette manière, on voit que x_n sera une solution du problème posé.

4.3 - PROCÉDE PRATIQUE DE CALCUL

4.31 - LEMME PRELIMINAIRE 1

- Soit E_p un sous-espace vectoriel de R^n engendré par la famille libre de vecteurs $y_1, \dots, y_p$ formant la matrice $Y(n, p)$ dans la base canonique de R^n .

- Soit $x_i, i \in I$ une famille dénombrable d'éléments de R^n formant la matrice $X (n \times |I|)$ dans la base canonique de R^n .

- La matrice $M = {}^t Y X$ est l'expression des projections orthogonales $\bar{x}_i$ des vecteurs $x_i, \forall i \in I$ sur E_p rapporté à la base $[y_1^*, \dots, y_p^*]$ duale de la base $[y_1, \dots, y_p]$.

DEMONSTRATION :

Soit $x_i \in R^n$ de matrice canonique X_i , on peut écrire $x_i = \bar{x}_i + x_i^-$, $x_i^- \in E_p^\perp$ orthocomplémentaire de E_p dans R^n .

Si E_p est rapporté à la base, $[y_1^* \dots y_p^*]$, duale de la base $[y_1 \dots y_p]$, on aura :

$$x_i = x_{1i} y_1^* + \dots + x_{pi} y_p^* + x_i^-$$

$$\text{d'où } \langle y_j, x_i \rangle = x_{ji}$$

$$\text{d'où } {}^t Y X_i = \begin{bmatrix} x_{1i} \\ \vdots \\ x_{pi} \end{bmatrix}$$

donc ${}^t Y X$ représente bien les projections orthogonales de $x_i, \forall i \in I$ sur E_p rapporté à la base $y_1^* \dots y_p^*$.

4,32 - ETUDE DU PROBLEME $\sup [\langle c, x \rangle \mid x \in D_p]$ - METHODE SEQUENTIELLE DE RESOLUTION

D_p est défini par $D_p = V_p \cap D$ avec :

$$V_p = [x \mid x = x_0 + \lambda_0 c + \lambda_1 a_1 + \dots + \lambda_{p-1} a_{p-1}]$$

$$D = \bigcap_{j \in J} H_j \quad H_j = [x \mid \langle a_j, x \rangle + b_j \geq 0]$$

$x \in D_p$ entraîne :

$$(1) \lambda_0 \langle a_j, c \rangle + \lambda_1 \langle a_j, a_1 \rangle + \dots + \lambda_{p-1} \langle a_j, a_{p-1} \rangle + u_j \geq 0 \quad \forall j \in J$$

$$\text{avec } u_j = \langle a_j, x_0 \rangle + b_j$$

La fonction économique s'écrit :

$$(2) \langle c, x \rangle = \langle c, x_0 \rangle + \lambda_0 \langle c, c \rangle + \dots + \lambda_{p-1} \langle c, a_{p-1} \rangle$$

Le problème revient à trouver $\lambda : (\lambda_0, \lambda_1, \dots, \lambda_{p-1})$ rendant maximum (2) compte tenu des contraintes (1).

D'après le théorème de Farkas, le dual de ce problème est :

$$\text{maximer } -u_0 = -\langle u, y \rangle = -\sum_{j \in J} u_j y_j$$

compte tenu des contraintes :

$$(3) \quad \left| \begin{array}{l} y_j \geq 0 \quad \forall j \in J \\ \sum_{j \in J} \langle a_j, c \rangle y_j = - \langle c, c \rangle \\ \sum_{j \in J} \langle a_j, a_1 \rangle y_j = - \langle c, a_1 \rangle \\ \sum_{j \in J} \langle a_j, a_{p-1} \rangle y_j = - \langle c, a_{p-1} \rangle \end{array} \right.$$

La matrice simplexe initiale $S_{(0)}^p$ correspondante écrite en disposant dans la première ligne les variables marginales, dans la première colonne les seconds membres est :

$$S_{(0)}^p = \begin{array}{|c|c|c|c|c|} \hline u_0 & u_1 & & u_k & \\ \hline - \langle c, c \rangle & \langle a_1, c \rangle & & \langle a_k, c \rangle & \\ - \langle c, a_1 \rangle & \langle a_1, a_1 \rangle & & \langle a_k, a_1 \rangle & \\ & \dots\dots\dots & & & \\ - \langle c, a_{p-1} \rangle & \langle a_1, a_{p-1} \rangle & & \langle a_k, a_{p-1} \rangle & \\ \hline \end{array}$$

Les variables marginales u_j sont positives; en effet :

$$x_0 \in \overset{0}{D} \Rightarrow u_j = \langle a_j, x_0 \rangle + b_j > 0 \quad \forall j \in J$$

- le second membre a au moins une composante négative $-\langle c, c \rangle$, il convient donc d'utiliser la méthode de résolution duale de la méthode simplexe.

- d'après le lemme précédent, la sous-matrice de $S_{(0)}^p$ réduite aux p dernières lignes représente les composantes des projections $-\bar{c}, \bar{a}_j, \forall j \in J$ des vecteurs $-c, a_j \forall j \in J$ sur E_p sous-espace vectoriel parallèle à V_p rapporté à la base $c^* a_1^* \dots a_{p-1}^*$, et comme $c \in E_p$ le système 3 peut s'écrire :

$$\left| \begin{array}{l} \sum_{j \in J} \bar{a}_j y_j = -c \\ y_j \geq 0 \quad j \in J \end{array} \right.$$

Supposons connaître x_{p-1} tel que $\langle c, x_{p-1} \rangle = \sup [\langle c, x \rangle \mid x \in D_{p-1}]$.

D'après l'hypothèse 2, x_{p-1} appartient à exactement $p-1$ plans P_j . Soit $J_0 \subset J$ l'ensemble des j tels que $x_{p-1} \in P_j$. On a $u_j = \langle a_j, x_{p-1} \rangle + b_j = 0 \quad \forall j \in J_0$

et, en particulier : $p-1 \in J_0$.

En mettant en évidence la partition $J_0 \cup \bar{J}_0$ de J et la partition $\{1 \dots p-1\} \cup \{p\}$ de l'ensemble des p dernières lignes de $S_{(0)}^p$, on écrira :

$$S_{(0)}^p = \begin{array}{c} \begin{array}{cc} & J_0 & \bar{J}_0 \\ \begin{array}{c} 0 \\ b_1 \\ b_2 \end{array} & \begin{array}{c} U_1 \\ A_1 \\ A_2 \end{array} & \begin{array}{c} U_2 \\ A_3 \\ A_4 \end{array} \end{array} = \begin{array}{c} \begin{array}{c} [1 \dots p-1] \\ [p] \end{array} \end{array} \begin{array}{c} S_{(0)}^{p-1} \\ 1^{p-1} \end{array}$$

S_0^{p-1} est la matrice simplexe initiale du problème $\sup [\langle c, x \rangle \mid x \in D_{p-1}]$ où E_{p-1} est rapporté à la base $[c^*, \alpha_1^*, \dots, \alpha_{p-2}^*]$ de E_{p-1} duale de la base $[c, \alpha_1, \dots, \alpha_{p-2}]$

Comme $x_{p-1} \in P_j, \forall j \in J_0$, les projections orthogonales $\bar{\alpha}_j$ des normales intérieures $\alpha_j, j \in J_0$ sur E_{p-1} , constituent la base de E_{p-1} optimale pour le problème :

$\sup [\langle c, x \rangle \mid x \in D_{p-1}]$, la matrice du changement de base :

$$[c^*, \alpha_1^*, \dots, \alpha_{p-2}^*] \rightarrow [\bar{\alpha}_j, j \in J_0]$$

est A_1 de telle sorte que la matrice simplexe $\hat{S}_{(J_0)}^{p-1}$ correspondant à la solution s'écrit :

$$\hat{S}_{(J_0)}^{p-1} = \begin{array}{c} \begin{array}{cc} & J_0 & J_0 \\ \begin{array}{c} -u^1 A_1^{-1} b_1 \\ A_1^{-1} b_1 \end{array} & \begin{array}{c} 0 \\ 1 \end{array} & \begin{array}{c} u^2 - u^1 A_1^{-1} A_3 \\ A_1^{-2} A_3 \end{array} \end{array} \\ J_0 \end{array}$$

et on aura :

$$-c = \sum_{j \in J_0} \lambda_j \bar{\alpha}_j \quad \lambda_j \geq 0$$

la matrice $(\lambda_j)_{j \in J_0}$ est $A_1^{-1} b_1$

La matrice $[0, u^2 - u^1 A_1^{-1} A_3]$ représente les

$$u_j = \langle \alpha_j, x_{p-1} \rangle + b_j \quad \forall j \in J.$$

Considérons la matrice :

$$N = \begin{vmatrix} A_1 & 0 \\ A_2 & 1 \end{vmatrix}$$

elle est non singulière car A_1^{-1} existe et représente la matrice du changement de base dans E_p

$$[c^*, \alpha_1^*, \dots, \alpha_{p-1}^*] \rightarrow [\bar{\alpha}(J_0) \cup \alpha_{p-1}^*]$$

En rapportant E_p à cette dernière base, on aura :

$$S_{(J_0)}^p = \begin{vmatrix} -u^1 A_1^{-1} b_1 & 0 & u^2 - u^1 A_1^{-1} A_3 \\ A_1^{-1} b_1 & 1 & A_1^{-1} A_3 \\ b_2 - A_2 A_1^{-1} b_1 & 0 & A_4 - A_2 A_1^{-1} A_3 \end{vmatrix} = \begin{vmatrix} \hat{S}_{(J_0)}^{p-1} \\ \hat{1}^{p-1} \end{vmatrix}$$

$$\text{Si on écrit } \hat{S}_{(J_0)}^{p-1} = \begin{vmatrix} U^{p-1} \\ V^{p-1} \end{vmatrix}$$

U^{p-1} étant la première ligne de $\hat{S}_{(J_0)}^{p-1}$, on aura encore :

$$\hat{1}^{p-1} = 1^{p-1} - A_2^{p-1} V^{p-1}$$

en posant, pour plus de généralité : $A_2 = A_2^{p-1}$.

NOTATION :

En écrivant la base $\bar{\alpha}(J_0) \cup \alpha_{p-1}^*$ de E_p sous la forme $[\bar{\alpha}_{j(1)} \dots \bar{\alpha}_{j(p-1)}] \cup \alpha_{p-1}^*$

La projection orthogonale $\bar{\alpha}_k$ de α_k sur E_{p-1} s'écrira :

$$\bar{\alpha}_k = x_{j(1),k} \bar{\alpha}_{j(1)} + \dots + x_{j(p-1),k} \bar{\alpha}_{j(p-1)} + y_{p-1,k} \alpha_{p-1}^*$$

EVENTUALITES :

a) Il est facile de voir (voir abandon des hypothèses 1 et 2) que $\hat{1}^{p-1}$ de composantes $(y_{p-1,0} \dots y_{p-1,i} \dots)$ n'est pas identiquement nulle.

b) $y_{p-1,0}$ est négatif.

En effet, si $y_{p-1,0} > 0$, on ne pourrait remplacer α_{p-1}^* par un vecteur $\bar{\alpha}_j$ $j \in J_0$ suivant la méthode duale.

Donc, $y_{p-1,0} \leq 0$ et il existe au moins un $j \in J_0$ tel que $y_{p-1,j} < 0$, sinon le système 3 serait vide et, comme $D_p \neq \emptyset$, on aurait, d'après le théorème de Farkas (conséquence 3) $\sup [\langle c, x \rangle \mid x \in D_p] = \infty$ contrairement au fait que D_p est borné.

$$\text{Soit donc } k \text{ tel que : } \frac{u_k}{y_{p-1,k}} = \sup_{j \in J - \left(\frac{u_j}{y_{p-1,j}} \right)}$$

$$J^- = [j \mid y_{p-1,j} < 0]$$

On peut remplacer dans la base $\bar{\alpha}_{J_0} \cup \alpha_{p-1}^*$ de E_p , α_{p-1}^* par $\bar{\alpha}_k$. On obtient ainsi un point extrême x_p' de D_p et, d'après l'hypothèse 2, on doit avoir $x_p' \neq x_{p-1}$ donc $y_{p-1,0} \neq 0$ et, finalement, $y_{p-1,0} < 0$.

EN RESUME :

On voit qu'il est très simple de former séquentiellement les solutions $x_1 \dots x_p \dots x_n$ des problèmes :

$$\sup [\langle c, x \rangle \mid x \in D_1], \dots, \sup [\langle c, x \rangle \mid x \in D_p], \dots, \sup [\langle c, x \rangle \mid x \in D_n = D]$$

a) Il suffit de partir de :

$$S_0^1 = \begin{vmatrix} u_j \\ y_{0j} \end{vmatrix} \quad j \in [0] \cup J$$

$$\text{Avec} \quad \begin{vmatrix} u_j = \langle \alpha_j, x_0 \rangle + b_j \\ y_{0j} = \langle \alpha_j, c \rangle \\ \alpha_0 = -c \end{vmatrix}$$

$$\text{A l'étape courante } p (p \leq n), \text{ on connaît } J_0, \hat{S}_{(J_0)}^{p-1} = \begin{vmatrix} U^{p-1} \\ V^{p-1} \end{vmatrix} \\ l^{p-1} \text{ et } A_2^{p-1}.$$

$$\text{On définit } S_{(J_0)}^p = \begin{vmatrix} \hat{S}_{(J_0)}^{p-1} \\ \hat{l}^{p-1} \end{vmatrix}$$

$$\text{avec } \hat{l}^{p-1} = l^{p-1} - A_2^{p-1} \hat{V}_{(J_0)}^{p-1}$$

Puis, on détermine dans la ligne $\hat{l}^{p-1}$ l'indice k tel que :

$$\frac{u^k}{y_{p-1,k}} = \sup_{j \in J^-} \left(\frac{u_j}{y_{p-1,j}} \right)$$

$$J^- = [j \mid y_{p-1,j} < 0]$$

α_{p-1}^* est remplacé par $\bar{\alpha}_k$.

Puis, en continuant suivant l'algorithme dual, on obtient une solution x_p du problème

$$\sup [\langle c, x \rangle \mid x \in D_p].$$

On connaît alors le nouvel ensemble J_0 tel que :

$$J_0 = [j \mid \langle \alpha_j, x_p \rangle + b_j = 0].$$

Dans J_0 , on affecte de l'indice p l'un des vecteurs α_j tels que $j \neq 1, \dots, p-1$.

Le processus se termine quand $p = n$.

INTERPRETATION PARTICULIERE DE LA MATRICE SIMPLEXE.

Soit $[\bar{\alpha}_{j(1)} \dots \bar{\alpha}_{j(p)}]$ une base quelconque duale admissible de E_p . La projection orthogonale $\bar{\alpha}_k$ de α_k sur E_p s'écrit :

$$\bar{\alpha}_k = \sum_{i=1}^p x_{j(i),k} \bar{\alpha}_{j(i)}$$

Soit $[\bar{\alpha}_{j(1)}^* \dots \bar{\alpha}_{j(p)}^*]$ la base de E_p duale de la base $[\bar{\alpha}_{j(1)} \dots \bar{\alpha}_{j(p)}]$.

On a alors :

$$x_{j(i),k} = \langle \bar{\alpha}_k, \bar{\alpha}_{j(i)}^* \rangle = \langle \alpha_k, \bar{\alpha}_{j(i)}^* \rangle$$

$\bar{\alpha}_{j(i)}^*$ est parallèle à l'arête de D_p contenue dans $\cap \Pi_{j(1)}, 1 \in [1 \dots p] - [i]$

Soit dans E_p l'opérateur $V_{j(i)}$ de projection sur la direction de cette arête.

Comme $c \in E_p$, on aura $V_{j(i)}(c) = \lambda \bar{\alpha}_{j(i)}^*$

λ sera déterminé par :

$$1 = \langle \bar{\alpha}_{j(i)}, \bar{\alpha}_{j(i)} \rangle = \frac{1}{\lambda} \langle \alpha_{j(i)}, V_{j(i)}(c) \rangle$$

par suite :

$$x_{j(i),k} = \langle \alpha_k, \bar{\alpha}_{j(i)}^* \rangle = \frac{\langle \alpha_k, V_{j(i)}(c) \rangle}{\langle \alpha_{j(i)}, V_{j(i)}(c) \rangle}$$

Cette interprétation sera utile pour la suite.

ABANDON DES HYPOTHESE 1 ET 2 :

a) Abandon de l'hypothèse 1

Si la famille $c \cup \alpha_{(J_0)}$ n'est pas en position générale dans R^n , α_{p-1} peut être combinaison linéaire de $c, \alpha, \dots, \alpha_{p-2}$ supposés linéairement indépendants. α_{p-1} appartient alors à E_{p-1} et $\hat{1}^{p-1}$ est identiquement nulle.

En effet, si $\alpha_{p-1} \notin E_{p-1}$, on peut écrire

$\alpha_{p-1} = \bar{\alpha}_{p-1} + \tilde{\alpha}_{p-1}$ avec $\bar{\alpha}_{p-1} \in E_{p-1}$ sous espace orthocomplémentaire de E_{p-1} dans E_p .

α_{p-1}^* est orthogonal à E_{p-1} donc constitue une base de $E_{p-1}^\perp$ et est tel que :

$$1 = \langle \alpha_{p-1}^*, \alpha_{p-1} \rangle = \langle \alpha_{p-1}^*, \bar{\alpha}_{p-1} \rangle \leq \| \alpha_{p-1}^* \| \| \bar{\alpha}_{p-1} \| \text{ donc } \| \alpha_{p-1}^* \| \geq \frac{1}{\| \bar{\alpha}_{p-1} \|}$$

Tout vecteur x de E_p s'écrit : $x = \bar{x} + \lambda \alpha_{p-1}^*$

avec $\bar{x} \in E_{p-1}$

$$\text{donc : } \| x \|^2 = \| \bar{x} \|^2 + \lambda^2 \| \alpha_{p-1}^* \|^2 \geq \| \bar{x} \|^2 + \frac{\lambda^2}{\| \bar{\alpha}_{p-1} \|^2}$$

Si le coefficient λ est différent de 0, la norme de tout vecteur $x \in E_p$ est non bornée quand $\bar{\alpha}_{p-1} \rightarrow 0$ donc quand $E_p \rightarrow E_{p-1}$, ce qui est absurde. Par suite, $\lambda = 0$. Un tel cas sera dit cas de dégénérescence majeure.

Le vecteur α affecté de l'indice $p-1$ ne convient pas pour définir E_p à partir de E_{p-1} . On choisira un autre vecteur $\alpha_j, j \in J_0$ qui sera affecté de l'indice d'ordre $p-1$. J_0 étant fini, on

arrivera en un nombre fini d'étapes, ou bien à construire une nouvelle variété V_p de dimension p contenant V_{p-1} , ou bien tous les vecteurs a_j , $j \in J_0$ appartiennent à E_{p-1} .

Théorème 27 :

Si $a_j \in E_{p-1}$ pour tout indice j appartenant à J_0 , alors :

$$\langle c, x_{p-1} \rangle = \sup [\langle c, x \rangle \mid x \in D_{p-1}] = \sup [\langle c, x \rangle \mid x \in D]$$

DEMONSTRATION :

En effet, le maximum de $\langle c, x \rangle$, $x \in D_{p-1}$ étant atteint en x_{p-1} , on a :

$$-c = \sum_{j \in J_0} c_j a_j \quad c_j \geq 0$$

Puisque le système $-c = \sum_{j \in J_0} l_j a_j$, $l_j \geq 0$ n'est pas vide, d'après le théorème de Farkas :

$$\langle a_j, x - x_{p-1} \rangle \geq 0 \quad \forall j \in J_0 \quad \text{entraîne} \quad \langle -c, x - x_{p-1} \rangle \geq 0.$$

Par suite, $\langle c, x \rangle \leq \langle c, x_{p-1} \rangle$ pour tout

$$x \in X = [x \mid \langle a_j, x \rangle + b_j \geq 0, j \in J_0]$$

comme $D \subset X$, on a bien $\langle c, x_{p-1} \rangle = \sup [\langle c, x \rangle \mid x \in D]$

Il pourra donc se faire qu'un élément solution du problème soit obtenu en moins de n étapes.

b) Abandon de l'hypothèse 2

Il peut y avoir plus de $p-1$ plans P_j $j \in J$ contenant x_{p-1} , c'est-à-dire $|J_0| > p-1$ donc plus de $p-1$ variables marginales u_j , $j \in J$ nulles. C'est un cas de « dégénérescence duale » que l'on résoud par les méthodes classiques.

Remarques. Choix du point de départ x_0 .

La méthode suppose $\overset{\circ}{D} \neq \emptyset$ et que l'on connaît $x_0 \in \overset{\circ}{D}$ ou, à la limite, que $x \in D$ (voir la remarque précédente : abandon de l'hypothèse 2).

Deux cas à envisager :

- le système $\langle a_j, x \rangle + b_j \geq 0$, $\forall j \in J$, a une solution immédiate : x_0 , qui peut être prise comme point de départ.

- le système précédent n'a pas de solution immédiate.

Le système

$$\left| \begin{array}{l} \langle a_j, x \rangle + b_j + \xi_j \geq 0 \quad j \in J \\ \xi_j \geq 0 \end{array} \right.$$

a toujours une solution immédiate ; donc, on peut toujours trouver un élément :

$$z_0 = [z_i = x_i \quad \forall i \in N, z_j = \xi_j \quad \forall j \in I]$$

appartenant à l'intérieur $\overset{\circ}{D}_1$ du polyèdre $D_1 \subset \mathbb{R}^{|I| + |N|}$ défini par :

$$\begin{cases} \langle a_j, x \rangle + b_j + \xi_j \geq 0 \\ \xi_j \geq 0 \end{cases} \quad j \in J$$

D_1 contient D et tout point extrême de D est point extrême de D_1 . Si l'on cherche à maximiser la forme $-\sum_{j \in I} \xi_j$ dans D_1 , deux éventualités peuvent se produire.

Ou bien à l'optimum il existe $j \in I$ tel que $\xi_j > 0$. Dans ce cas, D est visiblement vide.

Ou bien, à l'optimum $\xi_j = 0 \quad \forall j \in I$; dans ce cas, le système $\langle a_j, x \rangle + b_j \geq 0 \quad j \in J$ a une solution x_0 qui peut être prise comme point de départ.

4.4 - VARIANTE DE LA METHODE SEQUENTIELLE INTERIEURE

Au lieu de former la suite $x_0 \dots x_p \dots x_n$ où x_p maximise $\langle c, x \rangle$ dans $D_p = V_p \cap D$, on peut se contenter de former une suite $x_0 \dots x'_p \dots x'_n$ où x'_p est un point extrême de D_p tel que $\langle c, x'_p \rangle \geq \langle c, x'_{p-1} \rangle$. Ayant formé un point extrême x'_n de D , on atteindra x_n solution du problème posé en un nombre fini d'étapes.

La procédure la plus simple pour former une telle suite consiste à procéder de la manière suivante :

On supposera en premier lieu que les hypothèses 1 et 2 précédentes sont vérifiées.

a) V_1 étant défini par $V_1 = [x \mid x = x_0 + \lambda_0 c]$, le polyèdre $D_1 = V_1 \cap D$ est de dimension 1. x'_1 point extrême de D_1 appartiendra, d'après l'hypothèse 2, à un seul plan $P_{j(1)}$. On peut, sans restreindre la généralité, supposer que $j(1) = 1$.

b) La variété $V_2 = [x \mid x = x_0 + \lambda_0 c + \lambda_1 a_1]$ est, d'après l'hypothèse 1 de dimension 2. x'_2 , point extrême de D_2 tel que $\langle c, x'_2 \rangle \geq \langle c, x'_1 \rangle$, s'obtiendra en une seule itération d'après la méthode exposée en 4.32.

x'_2 appartiendra, d'après l'hypothèse 2, à 2 plans $P_{j(1)}$ et $P_{j(2)}$. On peut supposer $j(2) = 2$ sans restreindre la généralité.

c) Plus généralement, x'_p , $p \leq n-1$, point extrême de D_p appartient aux plans $P_1, P_2, \dots, P_{p-1}, P_p$. On définira $V_{p+1} = [x \mid x = x_0 + \lambda_0 c + \dots + \lambda_p a_p]$ qui est de dimension $p+1$ et, en une seule itération, on obtiendra x'_{p+1} point extrême de D_{p+1} tel que $\langle c, x'_{p+1} \rangle \geq \langle c, x'_p \rangle$ et qui, d'après l'hypothèse 2 appartient aux plans $P_1, P_2, \dots, P_p, P_{j(p+1)}$. On supposera que $j(p+1) = p+1$ sans restreindre la généralité.

En n étapes, on obtiendra ainsi x'_n qui est point extrême de $D_n = D$ tel que $\langle c, x'_n \rangle \geq \langle c, x'_{n-1} \rangle \geq \dots \geq \langle c, x'_1 \rangle \geq \langle c, x_0 \rangle$

L'abandon de l'hypothèse 2 conduit comme précédemment à la dégénérescence du problème. Il y aura donc plus de 1 candidat $P_{j(p+1)}$ à l'étape $p+1$ auquel on pourra donner l'indice $p+1$.

L'abandon de l'hypothèse 1 conduit à admettre que $\alpha_{p+1} \in E_p$. Dans ce cas :

- Ou bien x'_p est le maximum de $\langle c, x \rangle$ dans D_p et, d'après le théorème 27, le maximum de $\langle c, x \rangle$ dans D est atteint en x'_p .
- Ou bien x'_p n'est pas le maximum de $\langle c, x \rangle$ dans D_p et il est possible par itération, ou bien d'obtenir dans D_p un élément extrême x''_p tel que $\langle c, x''_p \rangle \geq \langle c, x'_p \rangle$ et appartenant à p plans $P_{j_1} \dots P_{j(p)}$ tels que $c, \alpha_{j(1)} \dots \alpha_{j(p)}$ soit une famille libre de $p+1$ vecteurs; dans ce cas, le processus pourra se poursuivre, ou bien un tel élément x''_p n'existe pas, et : $\sup [\langle c, x \rangle \mid x \in D_p] = \sup [\langle c, x \rangle \mid x \in D]$.

4,5 - Démontrons maintenant un théorème qui sera utile pour la méthode du gradient itéré.

Appelons x'_p le point extrême de D_p atteint suivant la variante de la méthode séquentielle en p étapes et x'_{p+1} le point extrême de D_{p+1} obtenu à l'étape suivante.

Théorème 28 :

Si V désigne l'opérateur de projection orthogonale défini dans E_{p+1} sur la direction de l'intersection I_p des plans $\Pi_1 \dots \Pi_p$ de V_{p+1} contenant x'_p , alors :

$$x'_{p+1} - x'_p = \lambda V(c), \quad \lambda \geq 0.$$

DEMONSTRATION :

En passant de la base $[\alpha_1, \dots, \alpha_p, \alpha_p^*]$ de E_{p+1} à la base $[\alpha_1, \dots, \alpha_p, \bar{\alpha}_{p+1}]$ ($p+1$ étant défini par $\frac{u_{p+1}}{y_{p,p+1}} = \sup_{j \in J^-} \left(\frac{u_j}{y_{p,j}} \right)$ $J^- = [j \mid y_{p,j} < 0]$)

On aura :

$$u'_j = u_j - u_{p+1} \frac{y_{p,j}}{y_{p,p+1}} \geq 0 \quad \forall j \in J$$

$$\text{Or : } u_j = \langle \alpha_j, x'_p \rangle + b_j$$

$$u'_j = \langle \alpha_j, x'_{p+1} \rangle + b_j$$

D'où, pour tout $j \in J$:

$$u'_j - u_j = \langle \alpha_j, x'_{p+1} - x'_p \rangle = - u_{p+1} \frac{y_{p,j}}{y_{p,p+1}} = - u_{p+1} \cdot x_{p+1,j}$$

Or, d'après la remarque sur l'interprétation de la matrice simplexe :

$$x_{p+1,j} = \frac{\langle \alpha_j, V(c) \rangle}{\langle \alpha_{p+1}, V(c) \rangle}$$

V étant l'opérateur de projection orthogonale sur l'arête de D_{p+1} contenue dans I_p .

Donc :

$$\langle \alpha_j, x'_{p+1} - x'_p \rangle = -u_{p+1} \frac{\langle \alpha_j, V(c) \rangle}{\langle \alpha_{p+1}, V(c) \rangle} \quad \forall j \in J$$

par suite :

$$x'_{p+1} - x'_p = - \frac{u_{p+1} V(c)}{\langle \alpha_{p+1}, V(c) \rangle} + y$$

$y \in \mathbb{R}^n$ étant orthogonal à tout α_j , $j \in J$ est nul.

$$\text{Posons : } \frac{-u_{p+1}}{\langle \alpha_{p+1}, V(c) \rangle} = \lambda$$

$$\text{On aura : } \langle c, x'_{p+1} \rangle - \langle c, x'_p \rangle = \lambda \langle c, V(c) \rangle$$

Comme : $\langle c, x'_{p+1} \rangle \geq \langle c, x'_p \rangle$ et $\langle c, V(c) \rangle \geq 0$ on aura $\lambda \geq 0$

On voit, de plus, que $\langle \alpha_{p+1}, V(c) \rangle < 0$. Le point extrême x'_{p+1} de D_{p+1} se déduit donc du point extrême x'_p de D_p par cheminement dans I_p suivant la projection orthogonale sur I_p du gradient c de la fonction économique.

Cette méthode a été présentée (d'une manière différente) par J. B. Rosen.

REMARQUE 1 :

Pour tout $j \in J$, on a $y_{p,j} = \theta \langle \alpha_j, V(c) \rangle$ θ étant une constante positive.

$$\text{En effet, } x_{p+1,j} = \frac{\langle \alpha_j, V(c) \rangle}{\langle \alpha_{p+1}, V(c) \rangle} = \frac{y_{p,j}}{y_{p,p+1}}$$

Donc : $y_{p,j} = \theta \langle \alpha_j, V(c) \rangle$ où θ est une constante non nulle. θ est positif car $y_{p,o} = \theta \langle -c, V(c) \rangle$ est négatif ou nul (voir le premier paragraphe « éventualités » de 4,32).

REMARQUE 2 :

Soit $J_0 = \{j \mid \langle \alpha_j, x'_p \rangle + b_j = 0\}$, on a $\langle \alpha_j, V(c) \rangle \geq 0 \quad \forall j \in J_0$.

En effet, pour $j \in [1, \dots, p]$, par définition de l'opérateur V , $\langle \alpha_j, V(c) \rangle = 0$.

Pour $j \in J_0 - [1, \dots, p]$ (cas de dégénérescence), on aura :

$$u'_j = -u_{p+1} \frac{\langle \alpha_j, V(c) \rangle}{\langle \alpha_{p+1}, V(c) \rangle} \quad \text{car } u_j = \langle \alpha_j, x'_p \rangle + b_j = 0 \text{ pour } j \in J_0$$

Or : $u'_j \geq 0$, $u_{p+1} \geq 0$, $\langle \alpha_{p+1}, V(c) \rangle < 0$, donc $\langle \alpha_j, V(c) \rangle \geq 0$.

REMARQUE 3 :

On peut généraliser, suivant la même méthode, la remarque 2. Si l'on passe du point extrême x'_p de D_p à un point extrême voisin x''_p tel que $\langle c, x''_p \rangle \geq \langle c, x'_p \rangle$ en remplaçant dans la base $\bar{\alpha}_{j(1)} \dots \bar{\alpha}_{j(p)}$ de E_p , le vecteur $\bar{\alpha}_{j(r)}$ tel que $x_{j(r),o} < 0$ par le vecteur $\bar{\alpha}_k$ tel que :

$$x_{j(r), k} < 0$$

$$\frac{u_k}{x_{j(r), k}} = \sup_{l \in J^-} \left(\frac{u_l}{x_{j(r), l}} \right) \quad J^- = [l \mid x_{j(r), l} < 0]$$

Alors, pour tout j tel que $\langle a_j, x'_p \rangle + b_j = 0$, on aura $\langle a_j, V_{j(r)}(c) \rangle \geq 0$.

★

★

★

CHAPITRE V

METHODE DU GRADIENT ITERE

5,1 - GENERALITES

5,11 - Si le polyèdre D des solutions admissibles du programme linéaire $\sup [\langle c, x \rangle \mid x \in D]$ comporte un très grand nombre de points extrêmes, la méthode simplexe, qui utilise un cheminement de point extrême en point extrême, nécessite, en général, un très grand nombre d'itérations.

Si l'intérieur $\overset{\circ}{D}$ de D n'est pas vide (on a vu que l'on pouvait toujours se ramener à ce cas) il semble intéressant d'utiliser une méthode de gradient mais on a vu au chapitre 3 que la méthode du double gradient de M. R. Frisch s'averait incertaine quant à l'obtention d'une solution optimale.

La méthode du gradient itéré va utiliser les avantages des méthodes de cheminement intérieur qui permettent d'augmenter rapidement la valeur de $\langle c, x \rangle$ dans les premières étapes et la sécurité des méthodes simpliciales qui permettent d'atteindre l'optimum en un nombre fini d'étapes.

5,12 - Soit $\succcurlyeq$ le préordre défini dans D

$y \succcurlyeq x$ si $\langle c, y \rangle \geq \langle c, x \rangle$ induisant la relation stricte $y \succ x$ si $\langle c, y \rangle > \langle c, x \rangle$. Soit X le sous ensemble des solutions du problème $\sup [\langle c, x \rangle \mid x \in D]$ supposé non vide.

Supposons que pour tout $x \in D - X$ supposé non vide on puisse former $y \in D$ tel que $y \succ x$. Alors la fonction $\langle c, x \rangle$ est strictement croissante suivant le filtre F d'éléments $[x \rightarrow [$ et sa limite suivant F est égale à sa borne supérieure $m = \sup [\langle c, x \rangle \mid x \in D]$.

PROCESSUS P_ϵ D'ORDRE FINI

- Soit ϵ un nombre strictement positif tel qu'il existe dans D au moins un x satisfaisant à :

$$m - \langle c, x \rangle \geq \epsilon$$

- Soit $X_\epsilon = [x \mid x \in D, m - \langle c, x \rangle < \epsilon]$

Si A est un sous-ensemble de D tel que $A \cap X \neq \emptyset$ on appellera processus P_ϵ d'ordre n une application f de A dans A telle que :

1. Si $x \in [A - A \cap X_\epsilon]$ il existe n fini tel que :

$$\left| \begin{array}{l} y = f^n(x) \in A \\ \langle c, y \rangle - \langle c, x \rangle \geq \epsilon \end{array} \right.$$

2. Si $x \in A \cap X_\epsilon$ il existe n fini tel que $y = f^n(x) \in X$

Il est immédiat qu'un processus P_ϵ d'ordre fini permet, à partir d'un point quelconque de A d'obtenir une solution du problème $\sup [\langle c, x \rangle \mid x \in D]$ en un nombre fini d'étapes.

EXEMPLE DE PROCESSUS P_ϵ D'ORDRE FINI :

Si on ordonne suivant la relation $\succcurlyeq$ les éléments $y \in \overset{\circ}{D}$ du profil de D , on obtient une suite finie, $y_1, \dots, y_p$.

$$\text{Soit : } \epsilon = \inf_{i, j \in [1 \dots p]} [|u_i - u_j| \mid u_i \neq u_j \mid u_i = \langle c, y_i \rangle]$$

avec ce choix de ϵ la méthode simplexe est un processus P_ϵ , en général d'ordre 1, défini dans $A = \overset{\circ}{D}$.

La méthode séquentielle intérieure est un processus P_ϵ d'ordre fini, défini dans tout D .

On va montrer que la méthode du gradient itéré est un processus P_ϵ d'ordre fini.

5,13 - LEMME PRELIMINAIRE

L'ensemble des éléments de la frontière $F(D)$ de D se divise en trois classes :

1. La classe supérieure $F(D)^+$ formée des éléments tels qu'il soit possible de les atteindre par un déplacement suivant c à partir d'un élément de l'intérieur de D .
2. La classe inférieure $F(D)^-$ formée des éléments tels qu'un déplacement suivant c permette d'atteindre un élément intérieur à D .
3. La classe complémentaire $F(D)^c$ qui est en général de mesure nulle dans $F(D)$.

Ce lemme est immédiat et se présente plutôt comme une définition des éléments supérieurs et inférieurs de D , car si $\overset{\circ}{D}$ n'est pas vide, à partir d'un point frontière de D ou bien c est dirigé vers l'extérieur de D ou bien vers l'intérieur de D , sauf sur le « contour apparent » de D suivant c qui est en général de mesure nulle dans $F(D)$.

5,14 - PRINCIPE DE LA METHODE DU GRADIENT ITERE

Supposons connu un élément $x_0 \in D$ et un procédé permettant de reconnaître qu'un élément de D est soit inférieur, soit supérieur, soit élément du contour apparent.

a. Si x_0 est un élément intérieur de D ou inférieur de D :

on peut construire un élément y supérieur de D $y_0 = x_0 + \lambda c$ $\lambda > 0$.

Trois éventualités peuvent se produire :

1. $y_0 \in X$ c'est une solution du problème.
2. $y_0 \notin X$ mais $\langle c, y_0 \rangle - \langle c, x_0 \rangle = \lambda \langle c, c \rangle \geq \epsilon$
Le processus est localement $P(\epsilon)$ d'ordre 1 et d'après le théorème 29 ci-après on peut construire $x_1 \in D$ tel que : $\langle c, x_1 \rangle = \langle c, y_0 \rangle$ x_1 joue le même rôle que x_0 et on est ramené au point a. précédent.

3. $y_0 \notin X$ et $\langle c, y_0 \rangle - \langle c, x_0 \rangle < \epsilon$

L'étape ayant permis de construire y_0 peut être considérée comme la première étape de la méthode séquentielle intérieure donc au bout d'un nombre fini d'étapes on aura construit :

$y \in F(D)$ tel que

ou bien $y \in X$ si $x_0 \in X$

ou bien $\langle c, y \rangle - \langle c, x_0 \rangle \geq \epsilon$ si $x_0 \notin X$

Trois cas sont à envisager dans cette dernière éventualité.

3,1 Si l'élément y est élément inférieur, il peut jouer le même rôle que x_0 du point a. précédent.

3,2 Si y est élément supérieur, il peut jouer le même rôle que l'élément y_0 du point a.2.

3,3 Si y est élément du contour apparent de D suivant c il conviendra de poursuivre suivant la méthode séquentielle intérieure.

b. Si x_0 est un élément supérieur :

on se trouve dans le cas décrit en $\alpha.2$.

c. Si x_0 est un élément du contour apparent de D suivant c :

on est ramené au cas $\alpha.3,3$. La méthode du gradient itéré est donc un processus P_ϵ d'ordre fini, défini dans D , et elle permet de former la solution du problème $\sup [\langle c, x \rangle \mid x \in D]$ en un nombre fini d'étapes.

REMARQUES :

1. L'élément ϵ est arbitraire. Si $\theta = \frac{\text{mesure de } X_\epsilon}{\text{mesure de } D}$ est faible, le cas $\alpha.2$ se produira fréquemment. Si, au contraire, θ est voisin de 1, le cas $\alpha.3$ sera le plus fréquent.
2. ϵ peut être modifié en cours de calcul. Il semble qu'il soit intéressant de prendre ϵ faible en début de calcul, et grand en fin de calcul.

5.2 - CARACTERISATION DES ELEMENTS SUPERIEURS ET DES ELEMENTS INFERIEURS

a. Caractérisation des éléments supérieurs :

D'après la définition, si x est élément supérieur de D , $(x \in F(D)^+)$, il existe $x_0 \in \overset{\circ}{D}$ tel que : $x = x_0 + \lambda c$ $\lambda > 0$

Soit P_j , $j \in J_1 \subset J$ les plans contenant x , on a :

$$\langle \alpha_j, x \rangle + b_j = 0$$

$$\langle \alpha_j, x_0 \rangle + b_j > 0$$

$$\text{d'où : } \langle \alpha_j, c \rangle < 0 \quad \forall j \in J_1$$

b. Caractérisation des éléments inférieurs

On voit, de la même manière, que si x est élément inférieur $(x \in F(D)^-)$ de D , tout plan P_j le contenant sera tel que $\langle \alpha_j, c \rangle > 0$.

5.3 - THEOREME 29

Si y_0 est élément supérieur de D ne maximant pas $\langle c, x \rangle$ dans D , si D est d'intérieur non vide, il est toujours possible de trouver $x_1 \in \overset{\circ}{D}$ tel que $\langle c, x_1 \rangle = \langle c, y_0 \rangle$

DEMONSTRATION :

On suppose tout d'abord que y_0 n'appartient pas à $\overset{\circ}{D}$. y_0 étant élément supérieur appartient aux plans P_j $j \in J_0 \subset J$, les normales intérieures α_j , $j \in J_0$ engendrent un sous espace vectoriel E_p de dimension $p \leq |J_0|$. Soit $I_0 \subset J_0$ tel que $[\alpha_j]$, $j \in I_0$, soit une famille libre engendrant E_p .

Premier cas :

$c \notin E_p$: on peut toujours supposer que I_0 est l'ensemble $[1 \dots p]$ tel que :

1. $c \cup \alpha(J_0)$ engendre un sous espace vectoriel E_{p+1} de dimension $p+1$.

2. Si E_{p+1} est rapporté à la base $[a_1' \dots a_p' a_p^*]$ (a_p^* vecteur dual de a_p dans la base $c \cup a(I_0)$), la matrice simplexe du problème $[\sup \langle c, x \rangle \mid x \in D_{p+1}]$,

$D_{p+1} = D \cap (y_0 + E_{p+1})$ à ses coefficients marginaux $u_j = \langle a_j, y_0 \rangle + b_j$, $j \in J$ non négatifs.

Avec les mêmes notations que pour le théorème 28, soit d la projection orthogonale de $V(c)$ sur E_c sous espace vectoriel orthocomplémentaire de c . $V(c) \neq 0$

on obtient aisément $d = V(c) - \frac{\langle c, V(c) \rangle}{\langle c, c \rangle} c$

Soit : $x_1 = y_0 + \mu d$, $\mu > 0$, alors pour μ assez petit $x_1 \in \overset{\circ}{D}$ et $\langle c, x_1 \rangle = \langle c, y_0 \rangle$

En effet, tout d'abord, $\langle c, x_1 \rangle = \langle c, y_0 \rangle + \mu \langle d, c \rangle = \langle c, y_0 \rangle$

Puis : pour tout $j \in J$

$$u'_j = \langle a_j, x_1 \rangle + b_j = \langle a_j, y_0 \rangle + b_j + \mu \langle a_j, d \rangle = u_j + \mu \delta_j$$

a. Pour $j \in J_0$, on a : $u'_j > 0$ quel que soit $\mu > 0$

En effet, $j \in J_0 \Rightarrow u_j = 0$ donc $u'_j = \mu \delta_j$

$$\text{et } \delta_j = \langle a_j, d \rangle = \langle a_j, V(c) \rangle - \frac{\langle c, V(c) \rangle}{\langle c, c \rangle} \langle c, a_j \rangle$$

D'après la remarque 2. du théorème 28 $\langle a_j, V(c) \rangle \geq 0$ et puisque y_0 est élément supérieur $\langle c, a_j \rangle < 0$ donc comme $\frac{\langle c, V(c) \rangle}{\langle c, c \rangle} > 0$ on a bien $\delta_j > 0$.

b. Pour $j \in J - J_0$, on a : $u'_j > 0$ pour $\mu < \inf_{j \in J^-} \left(\frac{-u_j}{\delta_j} \right)$

avec : $J^- = [j \mid j \in J - J_0, \delta_j < 0]$

en effet, $u_j > 0 \quad \forall j \in J - J_0$

Si $J^- = \emptyset$ alors $u'_j > 0$ pour tout $\mu > 0$

Deuxième cas :

$c \in E_p$: y_0 est alors élément extrême de $D_p = D \cap V_p$ avec $V_p = y_0 + E_p$
si $\langle c, y_0 \rangle = \sup [\langle c, x \rangle \mid x \in D_p]$ alors d'après le théorème 27 $y_0 \in X$ ensemble des solutions du problème.

Si $y_0 \notin X$ il existe un élément extrême y_1 voisin de y_0 tel que $\langle c, y_1 \rangle \geq \langle c, y_0 \rangle$ et d'après la remarque 3 du théorème 28, il existe un opérateur de projection orthogonale $V_{j(i)}$ tel que :

$$\text{si : } d_{j(i)} = V_{j(i)}(c) - \frac{\langle c, V_{j(i)}(c) \rangle}{\langle c, c \rangle} c$$

$x = y_0 + \mu d_{j(i)}$ appartient à l'intérieur de D (avec $\mu > 0$).

On le démontre comme précédemment.

On avait tout d'abord supposé que y_0 n'appartenait pas au profil $\overset{\circ}{D}$ de D . Si, maintenant $y_0 \in \overset{\circ}{D}$ on est ramené au deuxième cas ci-dessus.

5,4 - METHODE DE CALCUL RELATIVE A LA METHODE DU GRADIENT ITERE

Nous suivons le plan du paragraphe 5,14.

a. Passage de $x_0 \in \overset{\circ}{D} \cup F(D)^-$ à y_0 élément de $F(D)^+$

Soit $x_0 \in \overset{\circ}{D} \cup F(D)^-$ et ϵ une quantité positive finie; le passage de x_0 à y_0 peut être considéré comme la première étape de la méthode séquentielle intérieure (voir chapitre IV).

$$\text{On connaît } S_{(0)}^1 = \begin{vmatrix} u_j \\ y_{0j} \end{vmatrix} \quad j \in [0] \cup J$$

$$\text{avec } \begin{vmatrix} u_j & = & \langle a_j, x_0 \rangle + b_j \\ y_{0j} & = & \langle a_j, c \rangle \\ a_0 & = & -c \end{vmatrix}$$

Sans restreindre la généralité on peut supposer que :

$$\frac{u_1}{y_{01}} = \sup_{j \in J^-} \left(\frac{u_j}{y_{0j}} \right) \quad \text{avec } J^- = [j \mid y_{0j} < 0]$$

($J^- \neq \emptyset$ car D est borné et non vide).

c^* est remplacé par $\bar{a}_1$ dans la base de E_1 et on obtient $J_0, \hat{S}^1, l^1$ ainsi que

$$\langle c, y_0 \rangle - \langle c, x_0 \rangle = \frac{u_1}{y_{01}} u_0$$

(on rappelle que $\hat{S}^1$ est la matrice simplexe correspondant à l'optimum, l^1 est une matrice ligne formée ici des :

$$\langle a_1, a_j \rangle, j \in [0] \cup J,) \quad J_0 = [j \in J \mid \langle a_j, y_0 \rangle + b_j = 0]$$

On formera ensuite :

$$S_{(0)}^2 = \begin{vmatrix} \hat{S}_{(J_0)}^1 \\ \hat{l}^1 \end{vmatrix}$$

$$\hat{l}^1 = l^1 - A_2^{-1} \hat{S}_{(J_0)}^1$$

$$\text{ici } A_2^{-1} = \langle a_1, a_1 \rangle$$

Les trois éventualités à envisager sont les suivantes :

Eventualité a.1 : $y_0 \in X$. D'après le théorème 27 on sera dans ce cas si $\hat{l}^1$ est identiquement nulle.

Eventualité a.2 : $y_0 \in X$. (c'est-à-dire $\hat{l}^1 \neq 0$) et $\langle c, y_0 \rangle - \langle c, x_0 \rangle = \frac{u_0 u_1}{y_{01}} \geq \epsilon$.

Dans ce cas il conviendra de construire $x_1 \in \overset{\circ}{D}$ tel que : $\langle c, x_1 \rangle = \langle c, x_0 \rangle$.

REMARQUE IMPORTANTE

En général le sous-ensemble J_0 sera réduit à un seul élément car si x_0 est pris au hasard dans $\overset{\circ}{D} \cup F(D)$ l'élément y_0 correspondant appartiendra presque certainement à un seul

plan d'appui car les éléments y_0 appartenant à plus d'un plan d'appui forment un ensemble de mesure nulle dans $F(D)$.

Nous examinerons successivement les deux cas où : $|I_0| = 1$ et $|I_0| > 1$.

α.2,1 - *Premier cas* : $|I_0| = 1$

On peut, en toute généralité, supposer que dans $S_{(0)}^2$

$$\frac{u_2}{y_{1,2}} = \sup_{j \in J^-} \left(\frac{u_j}{y_{1j}} \right) \quad J^- = [j \in J \mid y_{1j} < 0]$$

(en effet, y_{11} est nul et $J^- \neq \emptyset$ car $\langle c, x \rangle$ est bornée dans D).

D'après le théorème 29 on obtiendra un élément $x_1 \in \hat{D}$:

$$x_1 = y_0 + \mu d$$

tel que :

$$\begin{cases} \langle c, x_1 \rangle = \langle c, y_0 \rangle \\ u'_j = \langle a_j, x_1 \rangle + b_j = u_j + \mu \delta_j \end{cases}$$

avec :

$$\delta_j = \langle a_j, V(c) \rangle - \frac{\langle c, V(c) \rangle}{\langle c, c \rangle} \quad \langle c, a_j \rangle$$

Or, d'après la remarque 1 du théorème 28 on a :

$$\langle a_j, V(c) \rangle = \frac{1}{\theta} y_{1j} \quad \text{avec : } \theta > 0$$

Si on pose par définition :

$$x_{0j} = - \langle a_j, c \rangle$$

on aura :

$$\delta'_j = \theta \delta_j = [y_{1j} - y_{1,0} \frac{x_{0j}}{x_{00}}]$$

et x_1 sera intérieur à D pour tout μ tel que :

$$0 < \mu < \hat{\mu} = \inf \left[\left(\frac{-u_j}{\delta'_j} \right) \mid \delta'_j < 0 \right]$$

il semble avantageux de choisir $\mu = \frac{\hat{\mu}}{2}$, c'est-à-dire x_1 « voisin d'une position centrale » dans D .

Pour la suite du calcul x_1 se comporte comme l'élément x_0 précédent. On est ramené au point α.

α.2,2 - *Deuxième cas* : $|I_0| > 1$

L'étape ayant permis de définir y_0 est considérée comme la première étape de la méthode séquentielle intérieure.

On continuera suivant cette méthode jusqu'à ce que l'on puisse définir une augmentation de la forme $\langle c, x \rangle$.

Soit x' le point extrême de D_p tel que $\langle c, x'_p \rangle \geq \langle c, y_0 \rangle$ (ceci implique entre autres que y_0 est un élément extrême optimum pour $D_1, D_2, \dots, D_{p-1}$).

On sera alors en présence de la matrice :

$$S_{(J_0)}^p = \begin{vmatrix} \hat{S}_{(J_0)}^{p-1} \\ \hat{1}^{p-1} \end{vmatrix}$$

On voit immédiatement « mutatis mutandis » ce qui a été dit dans le premier cas précédent, que l'élément $x_1 = y_0 + \mu d$ appartient à D pour : $0 < \mu < \hat{\mu} = \inf \left[\left(\frac{-u_j}{\delta'_{j,j}} \right) \mid \delta'_{j,j} < 0 \right]$

$$\text{avec : } \delta'_{j,j} = \left[y_{p-1,j} - y_{p-1,0} \frac{x_{0j}}{x_{00}} \right] \quad \forall j \in J$$

et on aura :

$$u'_j = u_j + \mu \delta'_{j,j}, \quad j \in J.$$

Cet élément x_1 se comporte pour la suite du calcul comme l'élément x_0 précédent. On est ramené au point α .

Eventualité a.3 : $y_0 \notin X$ et $\langle c, y_0 \rangle - \langle c, x_0 \rangle < \epsilon$

On continuera suivant la méthode séquentielle intérieure jusqu'à l'étape p , $p \leq n$, où on aura trouvé un point extrême x'_p de D_p tel que : $\langle c, x'_p \rangle - \langle c, y_0 \rangle \geq \epsilon$

On pourra alors, si on a de bonnes raisons de penser qu'on est au voisinage de l'optimum, continuer selon la méthode séquentielle intérieure; sinon on cherchera un élément x_1 intérieur à D pour se ramener au cas α . Pour ce faire :

Supposons donc être en présence de la matrice simplexe $S_{(J'_0)}^p$ (qui a été déduite de $S_{(J_0)}^p$ par changement de base dans E_p).

- Si x'_p ne maxime pas $\langle c, x \rangle$ dans D_p : il existe dans la matrice $S_{(J'_0)}^p$ un $x_{j(i),0}$ négatif et un indice k tel que :

$$\frac{u_k}{x_{j(i),k}} = \sup \left[\left(\frac{u_l}{x_{j(i),l}} \right) \mid x_{j(i),l} < 0 \right]$$

D'après la remarque 3 du théorème 28 et le théorème 29 on voit comme précédemment que l'élément $x_1 = x'_p + \mu d_{j(i)}$ est intérieur à D pour :

$$0 < \mu < \hat{\mu} = \inf \left[\left(\frac{-u_l}{\delta'_{j(i),l}} \right) \mid \delta'_{j(i),l} < 0 \right]$$

$$\text{avec : } \delta'_{j(i),l} = \left[x_{j(i),l} - x_{j(i),k} \frac{x_{0l}}{x_{00}} \right]$$

et on aura :

$$u'_1 = \langle a_1, x_1 \rangle + b_1 = u_1 + \mu \delta'_{j(i),1}$$

- Si x'_p maxime $\langle c, x \rangle$ dans D_p : on formera la matrice simplexe $S_{(J_0)}^{p+1}$ relative au problème $\sup [\langle c, x \rangle \mid x \in D_{p+1}]$ et on se trouve ramené au cas $\alpha, 2, 2$.

b. Si x_0 est élément supérieur ($x_0 \in F(D)^+$)

(Il suffira de regarder les signes de x_{0j} pour $j \in J_0$).

On se retrouve dans le cas décrit à l'éventualité a.2.

c. Si x_0 appartient à $F(D)^c$

On se retrouve dans le cas décrit à l'éventualité a.3.

★

★

★

CHAPITRE VI

GENERALISATION DE LA METHODE DU GRADIENT ITERE A UNE CLASSE DE PROGRAMME NON LINEAIRE

La méthode du gradient itéré peut se généraliser aux problèmes du type suivant :

maximer $f(x)$

sous les conditions $g(x) \geq 0$

où $f(x)$ est une fonction numérique concave et $g(x)$ une application de R^n dans R^m de composantes $g_i(x)$, $i \in [1, \dots, m]$ concaves, et de classe C^2 (c'est-à-dire de dérivées seconde continues) et telles qu'il existe un $x = x_0$ satisfaisant à $g_i(x_0) > 0$.

1 - GENERALITES

1.1. Il est facile de voir que, dans ces conditions, le domaine D des solutions possibles est un convexe de R^n , d'intérieur $\overset{\circ}{D}$ non vide. On supposera que D est borné, ce qui est toujours possible à réaliser en introduisant au besoin les contraintes $a_i \leq x_i \leq b_i \quad \forall i \in [1, \dots, n]$.

Le problème posé est équivalent au problème maximiser z sous les conditions :

$$\begin{cases} g(x) \geq 0 \\ \varphi(x, z) = f(x) - z \geq 0 \end{cases}$$

En effet, soit $\bar{D}$ le domaine de R^{n+1} formé des couples (x, z) satisfaisant aux conditions précédentes. $\bar{D}$ est un convexe d'intérieur non vide et $D \subset \bar{D}$.

Notons $[x, z(x)]$ les éléments de $\bar{D}$ tels que $z = f(x)$

Pour tout $(x, z) \in \bar{D}$, on aura : $z \leq f(x) = z(x) \leq z(x_0) = f(x_0)$

x_0 étant tel que $z(x_0) = \sup [z \mid (x, z) \in \bar{D}]$ ($x_0 \in D$ car D est compact).

Donc, pour tout $x \in D$, on aura $f(x) \leq f(x_0)$

Donc, comme $f(x) - z$ est une fonction concave, le problème posé se ramène en toute généralité à l'étude du problème :

maximer la forme linéaire $\langle c, x \rangle$

sous la condition $g(x) \geq 0$ où $g(x)$ est une application de R^n dans R^m concave, de classe C^2 et telle qu'il existe x_0 satisfaisant à $g(x_0) > 0$.

Théorème 30 :

Le domaine X des solutions du problème $\sup [\langle c, x \rangle \mid g(x) \geq 0]$ est un sous ensemble de la frontière $F(D)$ de D .

En effet, supposons que $x_0 \in \overset{\circ}{D}$, (intérieur de D), soit solution du problème, il existe alors $y = x_0 + \lambda c$, $\lambda > 0$ tel que $y \in F(D)$ puisque D est borné et $\langle c, y \rangle = \langle c, x_0 \rangle + \lambda \langle c, c \rangle$ donc $\langle c, y \rangle > \langle c, x_0 \rangle$ ce qui est impossible.

Théorème 31 :

CONDITIONS DE KUHN ET TUCKER :

Enoncé particulier :

1. $f(x)$ est une fonction numérique concave (de classe C^1), $g(x)$ une application concave

de R^n dans R^m de classe C^1 une condition nécessaire et suffisante pour que $\hat{x}$ maximise $f(x)$ sous les conditions $g(x) \geq 0$ est qu'il existe dans R^m un vecteur u tel que :

$$(1) \quad u \geq 0$$

$$(2) \quad \langle u, g(\hat{x}) \rangle = 0$$

$$(3) \quad \left(\frac{dg}{dx} \right)_{\hat{x}} \cdot u + \left(\frac{df}{dx} \right)_{\hat{x}} = 0$$

REMARQUES

1. D'autres énoncés plus généraux existent, en particulier celui de Slater qui impose seulement à f et g d'être continues (c'est-à-dire de classe C^0). L'énoncé précédent, plus restrictif, s'applique puisqu'il a été supposé que f et g étaient de classe C^2 .
2. Les conditions (2) sont appelées conditions de dualité ou conditions d'exclusion. En effet :
 $g_i(\hat{x}) > 0 \Rightarrow u_i = 0$
 $u_i > 0 \Rightarrow g_i(\hat{x}) = 0$
3. Les conditions (3) sont susceptibles d'une traduction géométrique intéressante. A l'optimum, l'opposé du gradient de la fonction économique est contenu dans le cône engendré par les gradients des contraintes $g(x_i)$ telles que $g_i(\hat{x}) = 0$.

2 - EXTENSION DE LA METHODE DU GRADIENT ITERE

2.1. La classification des éléments de $F(D)$ en éléments inférieurs, éléments supérieurs et éléments du contour apparent s'étend d'elle-même au cas où D n'est plus un polyèdre mais un convexe fermé borné. Il suffit de considérer les plans tangents à D au lieu des plans d'appui du polyèdre.

Soit $\epsilon \geq \eta > 0$, on va montrer que la méthode du gradient itéré est un processus P_ϵ défini dans tout D .

a. En effet, soit x_0 un élément intérieur de D ou inférieur de D .

On peut alors construire un élément supérieur de D , $y_0 = x_0 + \lambda c \quad \lambda > 0$

Trois éventualités peuvent se produire :

1. $y_0 \in X$, c'est une solution du problème. Il suffira, pour le vérifier, de s'assurer que $-c$ est une combinaison linéaire positive des gradients relatifs aux contraintes telles que $g_i(y_0) = 0$.

2. $y_0 \notin X$ mais $\langle c, y_0 \rangle - \langle c, x_0 \rangle = \lambda \langle c, c \rangle \geq \epsilon$. Le processus est localement P_ϵ et, en appliquant le théorème 29 aux plans tangents aux contraintes $g_i(x) \geq 0$ telles que $g_i(y_0) = 0$, on peut construire $x_1 \in D$ tel que $\langle c, x_1 \rangle = \langle c, y_0 \rangle$.

x_1 joue le même rôle que x_0 et on est ramené au point (a.) précédent.

3. $y_0 \notin X$ et $\langle c, y_0 \rangle - \langle c, x_0 \rangle < \epsilon$ (il importe de remarquer que la mesure de l'ensemble formé de tels éléments x_0 est de l'ordre de $K \cdot \epsilon$, avec K borné (K est inférieur à la mesure de $F(D)$)).

D'après la méthode du gradient projeté exposé ci-après, il est possible de trouver en un nombre fini d'étapes $y_1 \in F(D)$ tel que, ou bien $y_1 \in X$, ou bien tel que $\langle c, y_1 \rangle - \langle c, x_0 \rangle \geq \epsilon$.

Trois cas sont alors à envisager :

3,1 - Si y_1 est élément inférieur, il peut jouer le même rôle que l'élément x_0 du point (a) précédent.

3,2 - Si y_1 est élément supérieur, il peut jouer le même rôle que l'élément y_0 du point (a-2).

3,3 - Si y_1 est élément du contour apparent de D suivant c , on poursuivra suivant le processus résultant de la méthode du gradient projeté.

b. Si x_0 est un élément supérieur, on se trouve dans le cas décrit en (a-2).

c. Si y_0 est un élément du contour apparent de D , on est ramené au cas décrit en (a-3,3).

La méthode du gradient itéré est donc un processus P_ϵ d'ordre fini défini dans D tout entier, elle permet donc de former une solution du problème posé en un nombre fini d'étapes.

3 - LA METHODE DU GRADIENT PROJETE

La méthode du gradient projeté consiste à définir un cheminement sur la frontière de D suivant la projection du gradient de la fonction $\langle c, x \rangle$. Cette méthode n'est pas originale, elle a été développée par K. J. Arrow, L. Hurwicz, M. Solow, puis par J.B. Rosen.

Par contre, l'exposé ci-dessous de cette méthode est personnel.

Lemme 1.

Si y appartient à q variétés $g_i(x) = 0$ dont $p \leq n$ sont indépendantes, la projection $P(y)$ de c sur l'intersection Ω de ces variétés est localement lipchitzienne.

En effet, l'application $y \rightarrow P(y)$ est représentée dans la base canonique de R^n par la matrice :

$$P(y) = [I - B(y) ({}^t B(y) B(y))^{-1} {}^t B(y)] c = A(y) c$$

où $B(y)$ est la matrice jacobienne d'éléments

$$a_{ij} = \frac{\partial g_i}{\partial x_j}(y) \quad i \in [1, \dots, p] \quad j \in [1, \dots, n]$$

$({}^t B(y) B(y))^{-1}$ existe puisque les variétés $g_i(x) = 0 \quad i \in [1, \dots, p]$ sont indépendantes

Par ailleurs, puisque $B(y)$ est de rang p , il existe un voisinage $\omega \subset \Omega$ de y tel que p composantes y_i de y s'expriment continuellement en fonction des $n-p$ composantes restantes. On peut toujours supposer que les $n-p$ composantes y_i indépendantes forment le vecteur $y^1 = [y_1, \dots, y_{n-p}]$ d'un espace vectoriel E_{n-p} à $n-p$ dimensions. Si on désigne par $y^2 = \varphi(y^1)$ le vecteur formé par les p composantes non indépendantes de y , on aura :

$$P(y) = P(y^1, \varphi(y^1)) = \Pi(y^1)$$

Quand y parcourt $\omega \subset \Omega$, y^1 parcourt $\bar{\omega} \subset E_{n-p}$.

La différentielle $\Pi'(y^1)$ existe puisque les fonctions g_i sont de classe C^2 et si A désigne une boule fermée de E_{n-p} , de centre y^1 et contenue dans $\bar{\omega}$, en appelant M une borne de $\|\Pi'(y^1)\|$ dans A la formule des accroissements finis permet d'écrire :

$$\|\Pi(y^1) - \Pi(\bar{y}^1)\| \leq M \|y^1 - \bar{y}^1\| \quad \text{soit :}$$

$$\|P(y) - P(\bar{y})\| \leq M \|y^1 - \bar{y}^1\|$$

A chaque point de Ω , on peut associer une boule fermée contenue dans Ω centrée en ce point ou $P(y)$ est lipschitzienne.

Lemme 2.

Le système différentiel :

$$I \quad \frac{dy^1}{dt} = \bar{\Pi}(y^1)$$

ou $\bar{\Pi}(y^1)$ est la restriction du vecteur $\Pi(y^1)$ à ses $n-p$ premières composantes, $p \leq n$, est équivalent au système

$$II \quad \begin{cases} \frac{dy}{dt} = P(y) \\ g_i(y) = 0 \quad i \in I_0 = [1, \dots, q] \end{cases}$$

En effet :

$\bar{\Pi}(y^1)$ projection de $\Pi(y^1) = P(y)$ sur l'espace vectoriel normé complet E_{n-p} est localement lipschitzienne puisque $P(y)$ l'est. Par suite, d'après le théorème de Cauchy Lipschitz, si U désigne un ouvert de $R.E_{n-p}$, quel que soit l'élément (t_0, y_0^1) de U il existe une solution maximale unique $y^1(t)$ au système I satisfaisant à : $y^1(t_0) = y_0^1$

$$\text{L'élément de } R^n : y(t) = \begin{cases} y^1(t) \\ y^2(t) = \varphi(y^1(t)) \end{cases}$$

est défini si φ existe, c'est-à-dire si $B(y(t))$ est de rang p , $y(t)$ vérifie alors identiquement le système II. En particulier, $y(t)$ décrit une trajectoire dans Ω quand t varie.

PROCESSUS PERMETTANT D'ATTEINDRE UNE SOLUTION DU PROBLEME

Sil'on ajoute au système II précédent les contraintes restantes $g_i(y) \geq 0 \quad i \in I_1$, avec $I_0 \cup I_1 = I$ ensemble des indices des contraintes $g_i(x) \geq 0$, l'élément y est astreint à rester dans le domaine fermé borné $\Omega \cap F(D)$. L'élément correspondant y^1 de E_{n-p} est alors astreint à appartenir au sous-ensemble Δ de E_{n-p} défini par :

$$g_i(y^1, \varphi(y^1)) \geq 0 \quad i \in I_1$$

DEUX EVENTUALITES SONT A ENVISAGER :

Eventualité 1.

Si pour toute valeur de t , $\bar{\Pi}(y^1(t)) \neq 0$ la frontière de Δ est alors atteinte au bout d'un temps $t_1 \leq \frac{L}{V}$ avec $V = \inf \|\Pi(y^1(t))\|$ et $L = \text{diamètre de } \Delta$. L'élément correspondant $y(t)$ de Ω atteint dans les mêmes conditions la frontière de $\Omega \cap F(D)$, c'est-à-dire qu'il existe un sous-ensemble J_0 de I_1 tel que :

$$g_i(y(t_1)) = 0 \quad i \in J_0 \cup I_0$$

On se trouve ramené au problème précédent en envisageant maintenant le système :

$$\begin{aligned} \frac{dy}{dt} &= P(y) \\ g_i(y) &= 0 \quad i \in I_0 \cup J_0 \\ g_i(y) &\geq 0 \quad i \in I_1 - J_0 \end{aligned}$$

prenant pour $t = 0$ la valeur initiale $y_1 = y(t_1)$.

Le rang de la nouvelle matrice jacobienne $B(y)$ est au moins égal à p .

L'élément $y(t)$ est donc astreint à se déplacer dans des variétés Ω de dimensions décroissantes. Il convient donc d'envisager l'éventualité 2.

Eventualité 2.

Si pour une valeur de $t = \bar{t}$ (finie ou non) $\bar{\Pi}(y^1(\bar{t})) = 0$ avec $B(y(\bar{t}))$ de rang $p \leq |I_0|$ alors $y(\bar{t}) = \bar{y}$ est, soit une solution du problème (c'est-à-dire $\bar{y} \in X$), soit tel que $\langle c, \bar{y} \rangle$ est stationnaire dans un voisinage A dans Ω de $\bar{y}$.

En effet, la correspondance $\bar{\Pi}(y^1) \rightarrow P(y)$ est bijective et alors $\bar{\Pi}(y^1) = 0 \Rightarrow P(y) = 0$.

Donc, ici $P(\bar{y}) = 0$, c'est-à-dire c appartient au sous-espace propre $V(\bar{y})$ de l'opérateur linéaire $P(\bar{y})$, correspondant à la valeur propre 0.

$V(\bar{y})$ est engendré par les vecteurs $\gamma_i(\bar{y}) = \left(\frac{dg_i}{dx} \right)_{x = \bar{y}}$,

$i \in [1, \dots, p]$, gradients au point $\bar{y}$ des applications numériques $g_i(x) = 0$.

Remarquons que si $|I_0| > p$ il y a d'autres vecteurs $\gamma_i(\bar{y})$ contenus dans $V(\bar{y})$.

Deux cas sont à envisager :

1. Il existe p gradients $\gamma_i(\bar{y})$ tels que : $-c = \sum_{i=1}^p \lambda_i \gamma_i(\bar{y})$ avec $\lambda_i \geq 0 \forall i, \sum_1^p \lambda_i \neq 0$.

On s'assure alors facilement que les conditions de Kuhn et Tucker sont vérifiées, donc $\bar{y} \in X$ domaine des solutions du problème $\sup [\langle c, x \rangle \mid g(x) \geq 0]$. La recherche de ces p gradients se ramène au programme linéaire immédiat :

maximer $\langle c, x \rangle$

compte tenu de $\langle \gamma_i(\bar{y}), x - \bar{y} \rangle \geq 0 \quad i \in I_0$

2. Il n'existe pas p gradients tels que $-c = \sum_1^p \lambda_i \gamma_i(\bar{y}), \lambda_i \geq 0 \forall i, \sum_1^p \lambda_i \neq 0$ alors $\langle c, \bar{y} \rangle$ est stationnaire au point $\bar{y}$.

En effet, désignons par $\bar{V}(\bar{y})$ le sous-espace vectoriel parallèle à la variété linéaire tangente à Ω en $\bar{y}$. On aura alors, puisque $\bar{V}(\bar{y})$ et $V(\bar{y})$ sont orthocomplémentaires :

$c = P(\bar{y}) + u(\bar{y})$ avec $u(\bar{y}) \in V(\bar{y})$

Si A désigne une boule fermée de centre $\bar{y}$ contenue dans Ω et de rayon $dr \geq 0$ et si $y' \in A$, on aura :

$\langle c, y' \rangle - \langle c, \bar{y} \rangle = \langle P(\bar{y}) + u(\bar{y}), y' - \bar{y} \rangle$

Comme $y' - \bar{y}$ peut être au deuxième ordre près considéré comme un élément de $\bar{V}(\bar{y})$ et comme $P(\bar{y}) = 0$ alors $\langle c, y' \rangle = \langle c, \bar{y} \rangle$

$\langle c, y \rangle$ est bien stationnaire dans A .

REMARQUE :

Ce cas se produira, en particulier, si le rang de $B(\bar{y})$ est égal à n .

De tels éléments $\bar{y}$ de D sont les homologues des points extrêmes des polyèdres. On les désignera sous le nom d'éléments H-extrêmes de D .

Revenons au cas général ou $p \leq n$. Il convient de trouver un processus permettant de continuer à augmenter la forme $\langle c, x \rangle$.

On peut toujours supposer que $[1, \dots, p]$ est la base du programme linéaire précédent tel que $\langle c, x \rangle$ peut être augmenté en remplaçant l'indice i par l'indice r n'appartenant pas à $[1, \dots, p]$, deux tels indices existent toujours sinon $\bar{y} \in X$.

On aura alors $-c = \sum_{j=1}^p \lambda_j \gamma_j(\bar{y})$ avec $\lambda_j < 0$ pour $j = i$

Soit $\gamma_j^*(\bar{y})$ $j \in [1, \dots, p]$ la base duale de la base $\gamma_j(\bar{y})$ de $V(\bar{y})$.

Si $P_i(\bar{y})$ désigne la projection orthogonale de c sur γ_i^* on aura :

$$P_i(\bar{y}) = \theta \gamma_i^* \text{ avec } \theta = \frac{\langle c, P_i(\bar{y}) \rangle}{-\lambda_i}$$

Comme $\langle c, P_i(\bar{y}) \rangle$ est la forme semi-définie positive

$${}^t c [I - U(\bar{y}) [{}^t U(\bar{y}) U(\bar{y})]^{-1} {}^t U(\bar{y})] c$$

avec $U(\bar{y}) =$ matrice jacobienne d'éléments $\left(\frac{\partial g_j}{\partial x_k} \right)_{\bar{y}}$

$$j \in [1, \dots, p] - i, \quad k \in [1, \dots, n]$$

on aura $\theta > 0$

$$\text{par suite } g_j[\bar{y} + dt P_i(\bar{y})] = \theta \langle \gamma_i^*(\bar{y}), \gamma_j(\bar{y}) \rangle dt = \theta \delta_{ij} dt \quad j \in I_0$$

(δ_{ij} étant les symboles de Kronecker).

On peut donc revenir à la forme initiale du système différentiel

$$\begin{aligned} \frac{dy}{dt} &= P(y) \\ g_j(y) &= 0 \quad j \in I_0 - [i] \\ g_j(y) &\geq 0 \quad j \in I_1 \cup [i] \end{aligned}$$

avec $P(y)$ projection orthogonale de c sur l'intersection des variétés $g_j(x) = 0$, $j \in I_0 - [i]$ dont on cherchera la solution maximale satisfaisant à $y(t_0) = \bar{y}$.

On est donc ramené au point (a) précédent.

A ce stade, la méthode permet, à partir d'un point quelconque y_0 de $F(D)$ de définir un cheminement dans $F(D)$ par arcs continus $(y_0, \dots, y_1)$, $(y_1, \dots, y_2)$ $(y_{k-1}, \dots, y_k)$ tel que chaque arc $(y_{k-1}, \dots, y_k)$ est contenu dans une partie fermée d'une variété superficielle V_k homéomorphe à un fermé Δ_k de E_{n-p_k} , de dimension $d_k = n - p_k$.

Si l'éventualité 1 se produit dans V_k , le cheminement se poursuit dans V_{k+1} avec $d_{k+1} \leq d_k$. Une telle succession d'éventualités (éventualités 1) étant finie puisqu'il existe au plus $2^{|I|} - 1$ variétés possibles d'intersection non vide avec $F(D)$, l'éventualité 2 se produira obligatoirement au bout d'un nombre fini d'étapes. C'est-à-dire qu'il existera k fini tel que dans V_k , on aura un élément y_k soit solution, soit point stationnaire.

Si, en généralisant la notion de « famille de vecteurs en position générale », on énonce :

La famille g_i de fonctions numériques est en position générale dans R^n si c ne peut être combinaison linéaire de moins de n vecteurs $y_i(x)$, les cas de stationnarité envisagés ne peuvent se produire qu'aux points H-extrême qui sont en nombre finis (inférieurs à $C \mid I \mid^n$) et le maximum de $\langle c, x \rangle$ dans D sera obtenu en un nombre fini d'étapes en un point H-extrême.

Dans le cas général, il n'est pas sûr que la famille $g_i, i \in I$ soit en position générale.

- a. Si les fonctions g_i sont strictement concaves, les éléments stationnaires sont en nombre finis. En effet, s'ils étaient en nombre infini, il y aurait au moins une variété V_k en contenant un nombre infini. Comme V_k est borné et fermé dans $F(D)$ compact, V_k est compact, donc la suite des points stationnaires dans V_k aurait au moins une valeur d'adhérence A_k , il existerait donc un voisinage ouvert de A_k dans V_k tel que $\langle c, y \rangle$ serait stationnaire dans V_k ce qui est contradictoire avec l'hypothèse de stricte concavité. L'optimum est donc atteint (pas nécessairement en un point H-extrême) en un nombre fini d'étapes.
- b. Si toutes les fonctions $g_i, i \in I$ ne sont pas strictement concaves, comme on a supposé que les g_i étaient de classe C_2 , on peut partitionner l'ensemble I en deux sous-ensembles I' et I'' tel que g_i est strictement concave pour $i \in I'$ et est une variété linéaire pour $i \in I''$.

Les intersections V_k déterminées par au moins une $g_i, i \in I'$ sont strictement concaves, donc le nombre de points stationnaires de tels V_k est fini. Les autres V_k (déterminés par des $g_i, i \in I''$ linéaires) sont linéaires, elles sont en nombre fini. Celles où $\langle c, y \rangle$ est stationnaire ne peuvent contenir plus d'un élément d'une trajectoire superficielle ayant pour origine un point y_0 quelconque de $F(D)$. Donc, l'optimum du problème posé sera atteint en un nombre fini d'étapes.

La mise en application pratique de la méthode du gradient projeté consiste à intégrer le système différentiel par une méthode de calcul numérique telle que, par exemple, la méthode de Rungé Kutta.

Il n'est pas dans notre propos de développer ce point ici.

Dans la méthode du gradient itéré, le recours à la méthode du gradient projeté est exceptionnel si l'on part d'un élément x_0 intérieur à D .*

En effet, elle ne sera utilisée que si le processus n'est pas localement P_ϵ , ϵ étant une quantité non nulle mais arbitrairement petite.

Comme pour la méthode du gradient itéré appliquée au programme linéaire, il sera exceptionnel également que l'élément de $F(D)$, $y_0 = x_0 + \lambda c$, $x_0 \in D$, $\lambda > 0$ vérifie plus d'une contrainte $g_i(x) = 0$ (ceci parce que l'intersection de plusieurs variétés $g_i(x) = 0$ est de mesure nulle dans $F(D)$ et qu'il y a un nombre fini de telles intersections possibles).

Enfin, cette méthode donne lieu à des calculs généralement simples.

* Ce qui est toujours possible en incluant D dans un convexe approprié d'intérieur non vide. Il suffira, comme dans le cas du programme linéaire, de résoudre le programme

$$\text{maximer } -M \sum_{i \in I} y_i + \langle c, x \rangle$$

$$\text{compte tenu de } \begin{aligned} g_i(x) + y_i &\geq 0 & i \in I \\ y_i &\geq 0 \end{aligned}$$

ou M est une constante arbitrairement grande.

DEUXIEME PARTIE

PROGRAMMATION LINEAIRE STOCHASTIQUE

CHAPITRE I

Dans la première partie, nous avons étudié le problème :

$$\sup [\langle c, x \rangle \mid x \in D]$$

où D est un polyèdre convexe fermé, généralement déterminé par $Ax \leq b, x \geq 0$, où (A, b, c) est une donnée. Dans cette deuxième partie nous examinerons ce que l'on peut dire du problème précédent quand (A, b, c) est une aléatoire qui suit une loi de probabilité L donnée.

Un tel problème sera dit programme linéaire stochastique. Si la loi de probabilité L n'est pas connue, on aura affaire à un programme linéaire incertain. L'étude de ce type de problème ne sera pas abordée ici (encore qu'il soit dans une certaine mesure possible de l'aborder comme problème stochastique paramétré).

1,1 - POINTS DE VUE POSSIBLES

Pour fixer les idées nous examinerons l'interprétation classique suivante du problème :

$$\sup [\langle c, x \rangle \mid Ax \leq b, x \geq 0]$$

$x = [x_j, j \in J]$ décrit le niveau cherché des activités $j \in J$ d'une entreprise E .

$c = [c_j, j \in J]$ décrit la marge bénéficiaire unitaire pour chaque activité $j \in J$.

$b = [b_i, i \in I]$ caractérise le niveau des moyens de production si $i \in I_1$, et les niveaux de ventes possibles des biens produits si $i \in I_2$, avec $I = I_1 \cup I_2$.

$A = [a_{ij}, i \in I, j \in J]$ décrit la consommation unitaire des moyens de production $i \in I_1$ par l'activité j au niveau unitaire, si $i \in I_2$ la quantité de biens i mise en vente par l'activité j au niveau 1.

L'entrepreneur pourra chercher s'il existe des niveaux d'activité x possibles quel que soit (A, b, c) suivant la loi L . Si de tels niveaux existent il pourra choisir celui (ou ceux) qui maximise l'espérance ou l'utilité de la marge bénéficiaire totale (point de vue 1) ou bien il pourra chercher quelle sera l'espérance mathématique de sa marge bénéficiaire maximum (point de vue 2).

L'entrepreneur pourra chercher, au cas où il peut adapter très rapidement et sans frais supplémentaires son entreprise à l'occurrence de (A, b, c) , quelle sera l'espérance mathématique de sa marge bénéficiaire maximum (point de vue 3), ou si l'adaptation nécessite des coûts supplémentaires, l'entrepreneur pourra chercher à fixer un niveau d'activité x tel qu'il maximise l'espérance mathématique de sa marge bénéficiaire compte tenu des coûts d'adaptation (point de vue 4).

A travers ces quatre points de vue, nous pouvons dégager deux points de vue principaux relatifs au critère et également deux points de vue principaux relativement au domaine des solutions admissibles.

POINT DE VUE RELATIF AU CRITERE

Le critère peut être :

. Soit la maximisation de l'espérance mathématique de la forme linéaire $\langle c, x \rangle$ ce sera le point

de vue de l'action car il permet de définir un comportement « moyen » en présence des événements aléatoires (A, b, c) suivant la loi L .

- Soit la définition de l'espérance mathématique du maximum de la forme $\langle c, x \rangle$ et plus généralement l'étude de la fonction de répartition ou de distribution du maximum de $\langle c, x \rangle$, ce sera le point de vue de la prévision.

POINT DE VUE RELATIF AU DOMAINE DES SOLUTIONS ADMISSIBLES

Outre l'étude de l'existence du domaine des solutions admissibles D deux points de vue peuvent être adoptés :

- Ou bien on envisage le domaine des solutions uniformément admissibles (éventuellement vide), c'est-à-dire le domaine des solutions admissibles quel que soit (A, b, c) suivant la loi L .
- Ou bien on envisage l'ensemble des domaines possibles correspondant aux éléments aléatoires (A, b, c) suivant la loi L .

Dans ce cas on dira que le domaine des solutions admissibles est flou.

REMARQUE :

Au même titre que l'on peut définir à partir de la loi L des lois de probabilité conditionnelles, des lois de probabilité marginales, on pourra définir des problèmes linéaires stochastiques conditionnels ou marginaux.

Par ailleurs, le type de la loi L , discrète ou continue, permettra de définir des problèmes stochastiques discrets ou continus.

1.2 - RELATIONS ENTRE PROBLEMES STOCHASTIQUES ET PROBLEMES PARAMETRÉS

1.2.1. Avant d'aborder l'étude générale des problèmes stochastiques nous examinerons quelques propriétés des programmes paramétriques qui seront utiles par la suite et mettront en évidence les liaisons existant entre programmes stochastiques et programmes paramétriques*.

Soit le problème : maximiser $\langle c, x \rangle$

compte tenu de $Ax \leq b$, $x \geq 0$, A étant une matrice réelle m, n .

x peut être considéré comme un élément d'un espace vectoriel E_n isomorphe à R^n .

c , est un élément de E_n^* dual de E_n .

b , est un élément d'un espace vectoriel E_m isomorphe à R^m

et A , est un élément d'un espace vectoriel E_{mn} isomorphe à $R^{m \cdot n}$.

CONVENTION D'ECRITURE :

1. Sauf exception, on désignera par la même écriture un élément x de E_p et sa matrice représentative dans la base canonique de R^p .
2. E_d étant l'espace vectoriel produit $E_d = E_n \times E_m \times E_{m \cdot n}$ isomorphe à R^d avec $d = (m+1)(n+1) - 1$, formé des éléments $A \in E_{mn}$, $b \in E_m$, $c \in E_n^*$, on notera indifféremment par $\theta(A, b, c)$ ou $\theta(x)$ l'image de l'élément $x = (A, b, c)$ par une application θ de E_d dans un espace vectoriel F .

* Une partie de ce chapitre est pour une certaine part empruntée à un exposé fait par Monsieur le Professeur Fortet, à l'AFCALTI-SOFRO Groupe MPE (5 juin 1961).

1.22 - Le problème $\sup [\langle c, x \rangle \mid Ax \leq b, x \geq 0]$ revient à définir le scalaire $\varphi = \sup [\langle c, x \rangle \mid x \in D]$ où D est le domaine des solutions admissibles $D = [x \mid Ax \leq b, x \geq 0]$.

Si (A, b, c) n'est plus un élément fixé dans E_d mais un élément susceptible de parcourir E_d , φ est une application de E_d dans R .

Nous allons chercher, en premier lieu, quand A et c sont donnés et b un élément de E_m à quelles conditions $\varphi(b)$ existe et est borné.

Pour que $\varphi(b)$ existe il est en premier lieu nécessaire que :

$$D(b) = [x \mid Ax \leq b, x \geq 0] \neq \emptyset$$

Soit $B \subset E_m$ l'ensemble des $b \in E_m$ tels que $D(b) \neq \emptyset$

Théorème 1.

B est un cône convexe non vide :

En effet, B est visiblement non vide.

si : $b \in B$, alors $\lambda b \in B \quad \lambda \geq 0$

si : $b_1, b_2 \in B$ il existe $x_1 \geq 0$ et $x_2 \geq 0$ tels que

$$Ax_1 \leq b_1$$

$$Ax_2 \leq b_2$$

$$\text{donc : } A(\lambda x_1 + (1-\lambda)x_2) \leq \lambda b_1 + (1-\lambda)b_2$$

$$\text{avec : } 0 \leq \lambda \leq 1$$

$$\text{donc il existe } x = \lambda x_1 + (1-\lambda)x_2 \in D(\lambda b_1 + (1-\lambda)b_2)$$

par suite $\lambda b_1 + (1-\lambda)b_2 \in B$, donc B est un cône convexe.

Théorème 2.

Pour que $b \in B$, il faut et il suffit que b puisse être mis sous la forme $b = b_0 + k$ avec $0 \leq k$, $b_0 \in B_0$, ou B_0 est le cône polyédral convexe de E_m engendré par la fermeture convexe des éléments de E_m représentés par les colonnes de A .

En effet, si $b \in B$ il existe $x_0 \geq 0$ tel que $Ax_0 \leq b$ ou $Ax_0 + k = b$ avec $k \geq 0$. Soit $b_0 = Ax_0$ alors $b_0 \in B_0$ et $b = b_0 + k$.

Réciproquement, si $b = b_0 + k$ avec $b_0 \in B_0$, $k \geq 0$ il existe $x_0 \geq 0$ tel que $Ax_0 = b_0$ et par suite $Ax_0 \leq b$ donc $b \in B$.

Théorème 3.

Si pour A donné $b \in B$, alors $D(b)$ est borné si et seulement si il n'existe pas $y \geq 0$, $y \neq 0$, tel que $Ay \leq 0$.

Si $b \in B$ alors $D(b) \neq \emptyset$. Si $D(b)$ n'est pas borné, il existe une suite x_n de points de E_n tels que :

$$Ax_n \leq b \quad x_n \geq 0, \quad \lim_{n \rightarrow \infty} \|x_n\| = \infty$$

la suite $y_n = \frac{x_n}{\|x_n\|}$ est contenue dans la frontière F de la boule $B(0,1)$ de E_n centrée en 0 et de rayon 1 . F étant compacte les x_n ont au moins une valeur d'adhérence $y \in F$ donc telle que :

$$y \geq 0, \quad y \neq 0, \quad \|y\| = 1 \quad \text{et} \quad Ay \leq 0$$

Réciproquement si A étant donné il existe $y \geq 0, y \neq 0$ tel que $Ay \leq 0$, alors $D(b)$ n'est pas borné pour tout $b \in B$.

En effet, soit $b \in B$, il existe au moins un $x_0 \in D(b)$ l'élément $x(\lambda) = x_0 + \lambda y, \lambda \geq 0$ satisfait à $Ax(\lambda) \leq b, x(\lambda) \geq 0$ pour tout $\lambda \geq 0$ donc $D(b)$ n'est pas borné.

CORROLAIRE :

Si $D(b)$ est borné quel que soit $b \in B$, c'est-à-dire s'il n'existe pas $x \geq 0, x \neq 0$ tel que $Ax \leq 0$ alors $\varphi(b, c)$ existe quel que soit $c \in E_n^*$.

Théorème 4.

Si A étant donné, $D(b)$ n'est pas borné quel que soit $b \in B$, $\varphi(b, c)$ est bornée si et seulement si $c \in C$ en appelant C le sous-ensemble de E_n^* tel que $D^*(c) \neq \emptyset$ avec :

$$D^*(c) = \{ y \mid y \in E_n^*, y \geq 0, yA \geq c \}$$

En effet, si $D^*(c) \neq \emptyset$ il existe $y \geq 0$ tel que $c \leq yA$ et on a quel que soit $x \in D(b)$: $\langle c, x \rangle \leq \langle y, Ax \rangle \leq \langle y, b \rangle$ et on ne peut avoir $\varphi(b, c) = \infty$.

Réciproquement, si $\varphi(b, c) < \infty$ on sait qu'il existe $y \geq 0$ tel que :

$$\langle y, b \rangle = \sup [\langle c, x \rangle \mid x \in D(b)] \text{ et satisfaisant à } yA \geq c, y \geq 0 \text{ donc } D^*(c) \neq \emptyset.$$

EN RESUME :

A étant donné $\varphi(b, c)$ existe sur l'ensemble produit

$$P = B \times C \subset E_m \times E_n^*$$

Si A est tel qu'il n'existe pas de solution aux conditions $Ax \leq 0, x \geq 0, x \neq 0, \varphi(b, c)$ est définie sur l'ensemble $B \times E_n^*$.

On voit aisément, comme précédemment, que C est un cône convexe non vide de E_n^* et que la condition nécessaire et suffisante pour que $c \in C$ est que c puisse se mettre sous la forme : $c = c_0 - 1$ avec $0 \leq 1 \in E_n^*, c_0 \in C_0$, où C_0 est le cône polyédral convexe engendré par la fermeture convexe des éléments de E_n^* représentés par les lignes de A .

Théorème 5.

$\varphi(b)$ est une fonction concave dans B et $\varphi(c)$ est une fonction convexe dans C .

$\alpha.$ est une fonction concave de b .

En effet, soit $b_1, b_2 \in B$ il existe $x_1 \in D(b_1)$ et $x_2 \in D(b_2)$ tels que :

$$Ax_1 \leq b_1, \quad x_1 \geq 0, \quad \langle c, x_1 \rangle = \varphi(b_1)$$

$$Ax_2 \leq b_2, \quad x_2 \geq 0, \quad \langle c, x_2 \rangle = \varphi(b_2)$$

Si λ appartient à $[0, 1]$, $D(\lambda b_1 + (1-\lambda) b_2)$ n'est pas vide car
 $x = \lambda x_1 + (1-\lambda) x_2 \in D(\lambda b_1 + (1-\lambda) b_2)$ et par suite :

$$\begin{aligned}\varphi(\lambda b_1 + (1-\lambda) b_2) &\geq \langle c, x \rangle = \lambda \langle c, x_1 \rangle + (1-\lambda) \langle c, x_2 \rangle \\ &= \lambda \varphi(b_1) + (1-\lambda) \varphi(b_2)\end{aligned}$$

b. φ est fonction convexe de c .

En effet, A étant fixé et b appartenant à B , soit $c_1, c_2 \in C$. Pour tout $x \in D(b)$ et pour tout $\lambda \in [0, \dots, 1]$.

$$\langle \lambda c_1 + (1-\lambda) c_2, x \rangle \leq \lambda \sup_{x \in D} \langle c_1, x \rangle + (1-\lambda) \sup_{x \in D} \langle c_2, x \rangle$$

donc :

$$\varphi(\lambda c_1 + (1-\lambda) c_2) \leq \lambda \varphi(c_1) + (1-\lambda) \varphi(c_2)$$

Théorème 6.

$\varphi(b)$ est continue à l'intérieur de B ainsi que $\varphi(c)$ à l'intérieur de C

a. En effet, soit b_0 un point intérieur à B et $\beta(b_0, \rho)$ une boule fermée de centre b_0 et de rayon $\rho < d(b_0, F(B))$, ou $d(b_0, F(B))$ est la distance de b_0 à la frontière $F(B)$ de B ($\rho > 0$ car b_0 est intérieur à B).

Soit b un point intérieur à $\beta(b_0, \rho)$ et u et v les points où le diamètre contenant b_0 et b rencontre la sphère $S(b_0, \rho)$ frontière de $\beta(b_0, \rho)$, on a :

$$1. b = b_0 + \lambda(u - b_0) \quad 0 \leq \lambda \leq 1$$

$$2. b = b_0 - \lambda(v - b_0)$$

$\varphi(b)$ étant concave on a d'après 1. :

$$\varphi(b) \geq \lambda \varphi(u) + (1-\lambda) \varphi(b_0)$$

$$\text{d'où : } \varphi(b) - \varphi(b_0) \geq \lambda(\varphi(u) - \varphi(b_0)) \geq -\lambda M$$

où M est la borne de $|\varphi(u) - \varphi(b_0)|$ quand u décrit S .

D'après 2. on a :

$$b_0 = \frac{b}{1+\lambda} + \frac{\lambda}{1+\lambda} v$$

d'où :

$$(1+\lambda) \varphi(b_0) \geq \varphi(b) + \lambda \varphi(v)$$

et :

$$\varphi(b) - \varphi(b_0) \leq \lambda(\varphi(v) - \varphi(b_0)) \leq \lambda M$$

et par suite :

$$|\varphi(b) - \varphi(b_0)| \leq \lambda M = \frac{M}{\rho} \|b - b_0\|$$

donc $\varphi(b)$ est continue en tout point b_0 intérieur à B (elle est même localement lipschitzienne).

b. On voit de même que $\varphi(c)$ est continue à l'intérieur de C .

1,3 - Supposons maintenant A non fixé dans $E_{m \cdot n}$

Théorème 7.

$\varphi(A, b, c)$ est définie dans $\Pi = \bigcup_{A \in E_{m \cdot n}} [A] \times P(A)$

Π est un sous-ensemble connexe de E_d

$$P(A) = P(\lambda, A) \quad \forall \lambda > 0 \quad \text{et} \quad \varphi(\lambda, A) = \frac{1}{\lambda} \varphi(A)$$

. En effet, si A n'est pas fixé dans $E_{m \cdot n}$ le sous-ensemble P précédent devient $P(A)$ et visiblement $P(A) \neq \emptyset$ pour tout A . $P(A)$ est connexe comme produit de deux connexes $B(A)$ et $C(A)$ (parce que convexes). Comme $(O_{E_m}, O_{E_n}) \in P(A)$ quel que soit A , il en résulte que Π est connexe.

. Si $(b, c) \in P(A)$ il existe $x_0 \in E_n$ tel que : $Ax_0 \leq b, x_0 \geq 0, \langle c, x_0 \rangle = \varphi(A) < \infty$ par suite il existe $y_0 = \frac{x_0}{\lambda}$ tel que $(\lambda A)y_0 \leq b, \lambda > 0, y_0 \geq 0$, et $\langle c, y_0 \rangle = \varphi(\lambda A) = \frac{1}{\lambda} \varphi(A)$ donc $(b, c) \in P(\lambda A)$ d'où $P(A) \subset P(\lambda A)$ on voit de même que $P(A) \supset P(\lambda A)$ d'où $P(A) = P(\lambda A)$

1,4 -

1,41 - STRUCTURES

Soit (A, b, c) un élément de Π et $\sup [\langle c, x \rangle \mid x \geq 0, Ax \leq b]$ le programme linéaire correspondant. Ce programme est équivalent au programme $\sup [\langle \bar{c}, z \rangle \mid \bar{A}z = b, z \geq 0]$ après introduction des variables d'écart ($\bar{A}$ est une matrice $(m, n + m)$ et $\bar{c}$, est une matrice $(n + m, 1)$).

A tout élément (A, b, c) de Π correspond un programme linéaire $\sup [\langle \bar{c}, z \rangle \mid \bar{A}z = b, z \geq 0]$ et un sous-ensemble J_0 , à m indices, de $J = [1, \dots, m + n]$ appelé base optimale, tel que

$$(\bar{A})_{J_0}^{-1} b \geq 0 \quad (\bar{c})_{J_0} \quad (\bar{A})_{J_0}^{-1} \quad \bar{A} - \bar{c} \geq 0$$

Soit N le sous-ensemble de Π où un élément (A, b, c) admet plus d'une base optimale (cas de dégénérescence), il est facile de voir que N est un fermé de $E_d \cap \Pi$ de mesure nulle.

Deux éléments (A_0, b_0, c_0) et (A_1, b_1, c_1) de $\Pi - N$ seront dits équivalents si les deux problèmes $\sup [\langle \bar{c}_0, z \rangle \mid \bar{A}_0 z = b_0, z \geq 0]$ et $\sup [\langle \bar{c}_1, z \rangle \mid \bar{A}_1 z = b_1, z \geq 0]$ admettent même bases optimales.

Cette relation est bien une relation d'équivalence et la classe d'un élément (A, b, c) sera appelée une structure. On notera cette structure S_{J_0} si J_0 est la base optimale correspondant à (A, b, c) . L'ensemble des structures forme donc une partition de $\Pi - N$.

1,42 - ETUDE SOMMAIRE DU PARAMETRAGE D'UN PROGRAMME LINEAIRE

Si $(A_0, b_0, c_0) \in S_{J_0}$, $(A_0 + A_1, b_0 + b_1, c_0 + c_1)$ appartient à S_{J_0} , si :

$$(A_0 + A_1)_{J_0}^{-1} (b_0 + b_1) \geq 0$$

$$(\bar{c}_0 + \bar{c}_1)_{J_0} (\bar{A}_0 + \bar{A}_1)_{J_0}^{-1} (\bar{A}_0 + \bar{A}_1) - (\bar{c}_0 + \bar{c}_1) \geq 0$$

et alors :

$$\varphi(A_0 + A_1, b_0 + b_1, c_0 + c_1) = (\bar{c}_0 + \bar{c}_1)_{J_0} (\bar{A}_0 + \bar{A}_1)_{J_0}^{-1} (b_1 + b_2)$$

si $\bar{A}_1$ est tel que $(\bar{A}_1)_{J_0} (\bar{A}_0)_{J_0}^{-1} (\bar{A}_1)_{J_0} = \lambda (\bar{A}_1)_{J_0}$ (λ scalaire) on obtient aisément :

$$(\bar{A}_0 + \bar{A}_1)_{J_0}^{-1} = (\bar{A}_0)_{J_0}^{-1} - \frac{(\bar{A}_0)_{J_0}^{-1} (\bar{A}_1)_{J_0} (\bar{A}_0)_{J_0}^{-1}}{1 + \lambda}$$

relation qui permet d'écrire les relations précédentes sans avoir à inverser la matrice $(\bar{A}_0 + \bar{A}_1)_{J_0}$ ce qui simplifie considérablement les études de paramétrage.

De tels cas de simplification se rencontrent en particulier dans les éventualités suivantes :

1. $\bar{A}_1 = k \bar{A}_0$ k étant un scalaire, alors $\lambda = k$
2. $(\bar{A}_1)_{J_0} = k e_i {}^t e_j$ k scalaire, e_i = matrice colonne d'éléments tous nuls, sauf le jème égal à 1, alors :
 $\lambda = k {}^t e_j (\bar{A}_0)_{J_0}^{-1} e_i$
ce cas se rencontre dans la paramétrisation d'un coefficient de la matrice.
3. $(\bar{A}_1)_{J_0} = b {}^t e_i$, avec b matrice colonne quelconque, alors : $\lambda = {}^t e_i (\bar{A}_0)_{J_0}^{-1} b$
ce cas se rencontre dans la paramétrisation d'une colonne de la matrice.
4. $(\bar{A}_1)_{J_0} = e_i {}^t b$, alors $\lambda = {}^t b (\bar{A}_0)_{J_0}^{-1} e_i$
ce cas se rencontre dans la paramétrisation d'une ligne de la matrice.
5. $(\bar{A}_1)_{J_0} = b_1 {}^t b_2$, b_1 et b_2 étant deux matrices colonnes quelconques,
alors $\lambda = {}^t b_2 (\bar{A}_0)_{J_0}^{-1} b_1$
6. $(\bar{A}_1)_{J_0} = k (\bar{A}_0)_{J_0} N$, k étant un scalaire, N une matrice idem potente, par exemple
 $N = I - M({}^t M M)^{-1} {}^t M$ ou M est une matrice quelconque de plein rang, alors $\lambda = k$.

Si $(A_0 + A_1, b_0 + b_1, c_0 + c_1) \notin S_{J_0}$ il est aisé de suivre ce qui se passe en introduisant l'élément $(A_0 + \theta A_1, b_0 + \theta b_1, c_0 + \theta c_1)$ et en faisant varier θ de 0 à 1 en opérant les changements de base qui s'imposent.

★

★

★

CHAPITRE II

2,1 - GENERALITES

Dans ce chapitre, nous examinerons ce que l'on peut dire de la fonction de répartition du maximum $\varphi(A, b, c)$ de la forme $\langle c, x \rangle$ dans le domaine $D = [x \mid Ax \leq b, x \geq 0]$ quand (A, b, c) est un élément aléatoire dans E_d .

Pour que cette recherche ait un sens, il est nécessaire que $\Pr [(A, b, c) \in \Pi] = 1$ ou Π est le sous-ensemble E_d où φ est *définie* et *bornée* (voir chapitre I).

En pratique, il est bien évident que l'élément (A, b, c) n'est défini subjectivement que dans un domaine borné Δ de E_d et le problème aura un sens si $\Delta \subset \Pi$.

Le choix d'une loi de probabilité L pour décrire la distribution de (A, b, c) impose que cette loi ait un support contenu dans Π . Cette nécessité exclu en général à priori toutes les distributions de probabilité à support non borné.

Néanmoins, si on choisit comme représentant le phénomène une loi de probabilité L telle que $\Pr [(A, b, c) \in \Pi] = 1 - \epsilon$, ϵ étant un nombre positif suffisamment petit, on dira qu'il y a adéquation (d'ordre ϵ) entre la loi L et le phénomène.

C'est sous ces conditions que nous envisagerons le problème.

Théorème 8.

A et b étant fixés dans E_{mn} et E_m , c étant un élément aléatoire suivant une loi L à support Δ borné dans E_n^* tels que $[A] \cdot [b] \cdot \Delta \subset \Pi$, $\varphi(c)$ désignant le maximum de la forme $\langle c, x \rangle$ dans le polyèdre $D = [x \mid Ax \leq b, x \geq 0]$, on a : $E(\varphi(c)) \geq \varphi(E(c))$.

En effet, $\varphi(c)$ est une fonction convexe continue à l'intérieur de $U = [A] \cdot [b] \cdot \Delta$ et bornée sur U domaine borné.

Théorème 9.

A et c étant fixé dans E_{mn} et E_n^* , b étant un élément aléatoire suivant une loi L à support K borné dans E_m tel que $[A] \cdot K \cdot [c] \subset \Pi$, $\varphi(b)$ désignant le maximum de la forme $\langle c, x \rangle$ dans le polyèdre $D = [x \mid Ax \leq b, x \geq 0]$, on a : $E(\varphi(b)) \leq \varphi(E(b))$.

En effet, $\varphi(b)$ est une fonction concave continue à l'intérieur de $V = [A] \cdot K \cdot [c]$, et bornée sur V domaine borné.

2,2 - MAJORATIONS

Supposons que l'élément (A, b, c) de E_d suive une loi discrète L à support borné Δ dans Π , on a un certain nombre de majorations utiles pour la suite.

Soit $I = [1, \dots, n]$ la suite des modalités de l'élément (A, b, c) suivant L .

Désignons par :

a_{ki}^i	l'élément courant de	A_i
b_k^i	l'élément courant de	b_i
c_i^i	l'élément courant de	c_i

soient :

$$A^+ \text{ la matrice d'éléments } a_{kl} = \sup_{i \in I} (a_{kl}^i)$$

$$A^- \text{ la matrice d'éléments } a_{kl} = \inf_{i \in I} (a_{kl}^i)$$

$$b^+ \text{ la matrice d'éléments } b_k = \sup_{i \in I} (b_k^i)$$

$$b^- \text{ la matrice d'éléments } b_k = \inf_{i \in I} (b_k^i)$$

$$c^+ \text{ la matrice d'éléments } c_l = \sup_{i \in I} (c_l^i)$$

$$c^- \text{ la matrice d'éléments } c_l = \inf_{i \in I} (c_l^i)$$

$$\text{Soient : } D^+ = [x \mid \bar{A}x \leq b^+, x \geq 0], \quad D^- = [x \mid A^+x \leq b^-, x \geq 0]$$

$$\text{et : } D_i = [x \mid A_i x \leq b_i, x \geq 0]$$

On a de façon évidente :

$$D^- \subset D_i \subset D^+$$

et, si l'on désigne par $\Psi(c, D)$ le maximum de la forme $\langle c, x \rangle$ dans D , on a :

$$\Psi(c^-, D^-) \leq \Psi(c_i, D^-) \leq \Psi(c_i, D_i) \leq \Psi(c_i, D^+) \leq \Psi(c^+, D^+)$$

ou en appelant $\varphi(A, b, c)$ le maximum de $\langle c, x \rangle$ dans le domaine $[x \mid Ax \leq b, x \geq 0]$, on a :

$$\varphi(A^+, b^-, c^-) \leq \varphi(A_i, b_i, c_i) \leq \varphi(A^-, b^+, c^+)$$

BORNES DU SUPPORT DE LA DISTRIBUTION DE

Si $y^- = (A^+, b^-, c^-)$, ainsi que $y^+ = (A^-, b^+, c^+)$ sont effectivement deux valeurs réalisables, $\varphi(y^-)$ et $\varphi(y^+)$ seront les bornes inférieures et supérieures du support de la distribution de $\varphi(y)$ dans R .

GENERALISATION

Plus généralement, si on définit la relation d'ordre $\geq$ dans l'ensemble des modalités de (A, b, c) suivant la loi L discrète ou non :

$$(A, b, c) \geq (A', b', c') \quad \text{si} \quad \left| \begin{array}{l} a_{ij} \geq a'_{ij} \\ b_i \geq b'_i \\ c_j \geq c'_j \end{array} \right.$$

et, si on choisit deux éléments (A_0, b_0, c_0) , (A_1, b_1, c_1) tels que :

$$(A_1, b_1, c_1) \geq (A_0, b_0, c_0), \text{ on aura :}$$

$$\Pr [\varphi(A_1, b_0, c_0) \leq \varphi(A, b, c) \leq \varphi(A_0, b_1, c_1)] \geq p$$

$$\text{avec } p = \Pr [(A_0, b_0, c_0) \leq (A, b, c) \leq (A_1, b_1, c_1)]$$

Théorème 10.

Si les k composantes aléatoires de (A, b, c) ($k \leq d$) sont indépendantes en probabilité, on aura, quelle que soit la loi de probabilité L :

$$\Pr [\varphi(A_0 + t\alpha, b_0 - t\beta, c_0 - t\gamma) \leq \varphi(A, b, c) \leq \varphi(A_0 - t\alpha, b_0 + t\beta, c_0 + t\gamma)] \geq p(t) \text{ avec } p(t) = \left(1 - \frac{1}{t^2}\right)^k$$

en appelant A_0, b_0, c_0 les éléments $E(A), E(b), E(c)$ et α, β, γ , les éléments de composantes $\sigma_{a_{ij}}, \sigma_{b_{ij}}, \sigma_{c_{ij}}$

DEMONSTRATION

En effet, pour chacune des k composantes aléatoires x_i de (A, b, c) , on a la relation de Bienaymé Tchebicheff :

$$\Pr [E(x_i) - t\sigma \leq x_i \leq E(x_i) + t\sigma] > 1 - \frac{1}{t^2}$$

donc :

$$\Pr [A - t\alpha, b_0 - t\beta, c_0 - t\gamma \leq A, b, c \leq A_0 + t\alpha, b_0 + t\beta, c_0 + t\gamma] > \left(1 - \frac{1}{t^2}\right)^k$$

REMARQUE

Si le type de loi est précisé, par exemple loi normale, la majoration sera meilleure ; on aura :

$$p(t) = (2F(t) - 1)^k$$

$$\text{avec : } F(t) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^t e^{-\frac{u^2}{2}} du$$

Il est relativement facile de calculer :

$$\Psi(t) = \varphi(A_0 + t\alpha, b_0 - t\beta, c_0 - t\gamma)$$

Il s'agit d'un programme linéaire paramétré, à un seul paramètre. On obtient ainsi :

$$\Pr [\Psi(-t) \leq \varphi(A, b, c) \leq \Psi(t)] \geq \left(1 - \frac{1}{t^2}\right)^k$$

Si les composantes de $x = (A, b, c)$ ne sont pas indépendantes en probabilité, il est en général difficile de calculer la probabilité p . Néanmoins :

Théorème 11.

Si x suit une loi normale N dans E_d , on a, avec les mêmes notations que pour le théorème 10, le résultat suivant :

$$p(t) = \Pr [\chi_d^2 \leq t^2]$$

avec χ_d^2 variable χ^2 classique à d degrés de liberté.

EN EFFET :

Soit $x = (A, b, c)$ distribué normalement dans E_d suivant la distribution :

$$f(x) = \frac{1}{(2\pi)^{\frac{d}{2}} \sqrt{\det M}} e^{-\frac{1}{2} {}^t(x-x_0) M^{-1} (x-x_0)}$$

avec M matrice des covariances $\mu_{ij} = E[(x_i - x_{0i})(x_j - x_{0j})]$.

La forme ${}^t_y M y = E\left[\sum_1^d y_i (x_i - x_{0i})\right]^2$ est définie positive, donc de même forme ${}^t_y M^{-1} y$.

Le sous-ensemble S de E_d tel que :

$$S = \{y \mid {}^t_y M^{-1} y = t^2\}$$

est un ellipsoïde dont l'équation tangentielle est :

$${}^t_u M u = \frac{1}{t^2}$$

Par suite, les plans $x_i = x_{0i} + \epsilon t \sqrt{{}^t e_i M e_i}$ sont tangents à S ($\epsilon = 1$ ou -1 , e_i = matrice colonne d'éléments tous nuls, sauf le $i^{\text{ème}}$ égal à 1).

En remarquant que $\sqrt{{}^t e_i M e_i} = \sigma_i$, si V_t désigne le pavé de E_d de base

$$[x_{0i} - t\sigma_i, x_{0i} + t\sigma_i], \text{ on a :}$$

$$\Pr[x \in V_t] \geq K \int_{Q(x) \leq t^2} e^{-\frac{1}{2} Q(x)} dx$$

$$\text{avec } K = \frac{1}{(2\pi)^{\frac{d}{2}} \sqrt{\det M}}$$

$$\text{et } Q(x) = {}^t(x - x_0) M^{-1} (x - x_0)$$

or M^{-1} étant symétrique est de la forme :

$$M^{-1} = {}^t U D U$$

où :

$$\left\{ \begin{array}{l} U \text{ est une matrice orthogonale} \\ D \text{ est une matrice diagonale d'éléments } \frac{1}{\theta_i} \text{ avec } \theta_i \text{ valeur propre de } M \text{ (donc } \theta_i > 0). \end{array} \right.$$

Par suite, en faisant le changement de variables $y = U(x - x_0)$, on aura :

$$\int_{Q(x) \leq t^2} e^{-\frac{Q(x)}{2}} dx = \int_{Q_1(y) \leq t^2} e^{-\frac{Q_1(y)}{2}} dy$$

$$\text{avec } Q_1(y) = {}^t_y D y = \sum \frac{y_i^2}{\theta_i}$$

En faisant le nouveau changement de variable $z_i = \frac{y_i}{\sqrt{\theta_i}}$, on a :

$$\int_{Q_1(y) \leq t^2} e^{-\frac{Q_1(y)}{2}} dy = \sqrt{\theta_1 \dots \theta_d} \int_{Q_2(z) \leq t^2} e^{-\frac{Q_2(z)}{2}} dz$$

$$\text{or : } \sqrt{\theta_1 \dots \theta_d} = \sqrt{\det M}$$

et $Q_2(z) = r^2$ est la sphère de centre 0 et de rayon r ; par suite :

$$\int_{Q_2(z)} e^{-\frac{Q_2(z)}{2}} dz = A_d \int_0^t e^{-\frac{r^2}{2}} r^{d-1} dr$$

avec A_d aire de la sphère unité de E_d .

$$\text{Comme } A_d = \frac{2\pi^{\frac{d}{2}}}{\Gamma(\frac{d}{2})}$$

et :

$$\int_0^t e^{-\frac{r^2}{2}} r^{d-1} dr = 2^{\frac{d-1}{2}} \int_0^{\frac{t^2}{2}} e^{-x} x^{\frac{d-1}{2}} dx = 2^{\frac{d-1}{2}} \Gamma\left(\frac{d}{2}, \frac{t^2}{2}\right)$$

il vient finalement :

$$K \int_{Q(x)} e^{-\frac{Q(x)}{2}} dx = \frac{\Gamma\left(\frac{d}{2}, \frac{t^2}{2}\right)}{\Gamma\left(\frac{d}{2}\right)}$$

Si on remarque que la variable $\chi_d^2 = \sum_1^d x_i^2$ (les x_i étant des variables normales réduites indépendantes) est telle que :

$$\Pr(\chi_d^2 \leq t^2) = \frac{\Gamma\left(\frac{d}{2}, \frac{t^2}{2}\right)}{\Gamma\left(\frac{d}{2}\right)}$$

on aura :

$$\begin{aligned} \Pr[(A_0 - t\alpha, b_0 - t\beta, c_0 - t\gamma) \leq (A, b, c) \leq (A_0 + t\alpha, b_0 + t\beta, c_0 + t\gamma)] \\ = \Pr(x \in V_t) \geq \Pr(\chi^2 \leq t^2) \end{aligned}$$

La loi de χ^2 a été tabulée pour $d \leq 30$, pour $d > 30$; on peut admettre que $\sqrt{2\chi^2} - \sqrt{2d-1}$ suit une loi normale réduite.

REMARQUE :

1. Le même type de raisonnement s'applique, « mutatis mutandis », si l'élément $x = (A, b, c)$ a seulement $k < d$ composantes aléatoires suivant une loi normale.

2. Si l'on adopte la majoration :

$$p(t) = \left(1 - \frac{1}{t^2}\right)^d$$

et si l'on veut que $p(t) = 1 - \alpha$, il conviendra de considérer un intervalle

$$[-t\alpha, t\alpha] \text{ avec } t\alpha = \left[1 - (1 - \alpha)^{\frac{1}{d}}\right]^{-1/2};$$

si α est faible, les valeurs de t seront de l'ordre de $\sqrt{\frac{d}{\alpha}}$.

Il est donc très vraisemblable qu'un tel type de majoration est utile pour des distributions de (A, b, c) assez concentrées autour de la droite Δt de E_d définie par :

$$\Delta t = [(A, b, c) \mid (A, b, c) = (A_0, b_0, c_0) + t(\alpha, \beta, \gamma)]$$

Si ce n'est pas le cas, en particulier si les variables sont indépendantes et si d est élevé, ce type de majoration n'est pas recommandé. Il est certainement plus réaliste de procéder de la façon suivante qui consiste à trouver une approximation de la loi de distribution de φ

Soit J_0 la base optimale correspondant à $y = (A_0, b_0, c_0)$ (l'indice 0 signifiant espérance mathématique).

On assimilera $\varphi(A, b, c)$ à la forme bilinéaire :

$(\bar{c})_{J_0} (\bar{A})_{J_0}^{-1} b$ ($\bar{A}$ et $\bar{c}$ se déduisant de A et c par introduction des coefficients relatifs aux variables d'écart).

$$\text{Alors : } \varphi(A, b, c) = (\bar{c}_0 + \delta \bar{c})_{J_0} (\bar{A}_0 + \delta \bar{A})_{J_0}^{-1} (b_0 + \delta b)$$

$$\text{avec : } \delta \bar{A} = \bar{A} - \bar{A}_0, \quad \delta b = b - b_0, \quad \delta \bar{c} = \bar{c} - \bar{c}_0$$

Si ces quantités sont petites, on assimilera au deuxième ordre près :

$$(\bar{A}_0 + \delta \bar{A})_{J_0}^{-1} \approx (\bar{A}_0)_{J_0}^{-1} - (\bar{A}_0)_{J_0}^{-1} (\delta \bar{A})_{J_0} (\bar{A}_0)_{J_0}^{-1}$$

Alors :

$$\begin{aligned} \varphi(A, b, c) &= \varphi(A_0, b_0, c_0) + (\delta \bar{c})_{J_0} (\bar{A}_0)_{J_0}^{-1} b_0 \\ &- (\bar{c})_{J_0} (\bar{A})_{J_0}^{-1} (\delta A)_{J_0} (\bar{A})_{J_0}^{-1} b_0 + (c)_{J_0} (\bar{A})_{J_0}^{-1} \delta b_0 \end{aligned}$$

Comme $z_0 = (\bar{A}_0)_{J_0}^{-1} b_0$ est la solution du problème

$\sup [\langle \bar{c}, z \rangle \mid \bar{A}_0 z = b_0, z_0 \geq 0]$ et $u_0 = \bar{c}_{J_0} (\bar{A}_0)_{J_0}^{-1}$ est la solution duale correspondante, on aura :

$$\varphi(A, b, c) = \varphi(A_0, b_0, c_0) + \langle \delta \bar{c}, z_0 \rangle - \langle u_0, (\delta \bar{A})_{J_0} z_0 \rangle + \langle u_0, \delta b \rangle$$

donc :

$$E \varphi(A, b, c) = \varphi(A_0, b_0, c_0) = \varphi(EA, Eb, Ec)$$

et, si les variables sont indépendantes, on aura :

$$\sigma^2 \varphi = \sum_{j \in J_0} z_{0j}^2 \sigma_{\bar{c}j}^2 + \sum_{i \in I_0} u_{0i}^2 \sigma_{b_i}^2 + \sum_{u \in I_0 \cdot J_0} u_{0i} z_{0j} \sigma_{a_{ij}}^2$$

(les composantes de u_0 sont indicées par $i \in I_0$)

et $\varphi(A, b, c)$ pourra être considérée comme distribuée normalement (par application du théorème « central limite ») autour de $\varphi(EA, Eb, Ec)$ avec l'écart type $\sigma \varphi$.

Bien entendu, il ne peut s'agir que d'une approximation dont on ne peut préciser la validité.

2,3 - CALCUL DIRECT - DETERMINATION DES FONCTIONS CARACTERISTIQUES

Si la connaissance de la loi de $\varphi(A, b, c)$ par les bornes inférieure et supérieure du support de la distribution, majorante de la loi de probabilité, et approximation normale semble insuffisante, on aura recours au calcul direct qui permet de calculer la fonction caractéristique de $\varphi(A, b, c)$.

Soit K l'ensemble des bases admissibles, c'est-à-dire telles que l'ensemble $[S_k] \quad k \in K$ des structures correspondantes forment une partition de $\Pi - N$.

Soit $y = (A, b, c) \in \Pi$ et $z(y)$ une solution maximale du problème

$\sup [\langle \bar{c}, z \rangle \mid \bar{A}z = b, z \geq 0]$ (si $y \in \Pi - N$, la correspondance $y \rightarrow z(y)$ est univoque)

Supposons que y est distribué suivant une loi L continue de support Δ borné et contenu dans Π .

Si $y \in S_{J_0}$, on a $z(y) = \begin{pmatrix} (\bar{A})_{J_0}^{-1} b \\ 0 \end{pmatrix}$ et, par suite, si u est un vecteur de E_{n+m} de com-

posantes non nulles, la fonction caractéristique de $z(y)$ sera :

$$\phi_z(u) = E \left(e^{i \langle u, z \rangle} \right) = \int_{\Pi - N} e^{i \langle u, z \rangle} f(y) dy = \int_{\Delta} e^{i \langle u, z(y) \rangle} f(y) dy$$

puisque $e^{i \langle u, z \rangle} f(y)$ est une fonction bornée dans $\Pi - N$ de support dans Δ et, puisque N est de mesure nulle (l'intégrale étant prise au sens de Lebesgue).

En tenant compte de l'expression de $z(y)$, on aura :

$$\phi_z(u) = \sum_{J_0 \in K} \int_{(A, b, c) \in S_{J_0} \cap \Delta} e^{i t(u)_{J_0} (\bar{A})_{J_0}^{-1} b} f(\bar{A}, b, \bar{c}) d\bar{A}, db, d\bar{c}$$

de même, la fonction caractéristique de $\varphi(y)$ sera :

$$\phi_\varphi(t) = \sum_{J_0 \in K} \int_{(A, b, c) \in S_{J_0} \cap \Delta} e^{i t(\bar{c})_{J_0} (\bar{A})_{J_0}^{-1} b} f(\bar{A}, b, \bar{c}) d\bar{A}, db, d\bar{c}$$

puisque pour $y \in S_{J_0}$, $\varphi(y) = (\bar{c})_{J_0} (\bar{A})_{J_0}^{-1} b$ est une fonction bilinéaire continue sur Δ .

On obtient de même $E(z)$ de composantes $E(z_i) \quad i \in [1, \dots, n]$

$$E(z_i) = \sum_{J_0 \in K} \int_{(A, b, c) \in S_{J_0} \cap \Delta} (\bar{A}_{J_0}^{-1} b)_i f(\bar{A}, b, \bar{c}) d\bar{A}, db, d\bar{c}$$

et :

$$E(\varphi) = \sum_{J_0 \in K} \int_{(A, b, c) \in S_{J_0} \cap \Delta} (\bar{c})_{J_0} (\bar{A})_{J_0}^{-1} b f(\bar{A}, b, \bar{c}) d\bar{A}, db, d\bar{c}$$

Si, sur le plan théorique aucune difficulté ne se pose pour définir ces fonctionnelles, il en va tout autrement sur le plan pratique ; en effet, S_{J_0} est défini dans E_d par les relations

$$(\bar{A})_{J_0}^{-1} b \geq 0 \quad (\bar{c})_{J_0} (\bar{A})_{J_0}^{-1} \bar{A} - \bar{c} \geq 0$$

et l'intégration d'une fonction, même simple, dans de tels domaines présente des difficultés importantes.

Si $y = (A, b, c)$ est distribué suivant une loi discrète, les intégrales précédentes sont remplacées par des sommations ; rien n'est changé sur le plan théorique, mais les difficultés sont moins grandes sur le plan pratique. Dans ce cas, on peut déterminer $E[\varphi(A, b, c)]$ relativement facilement.

PREMIERE METHODE

Si l'ensemble des modalités de $y = (A, b, c)$ est fini avec $\Pr[(A, b, c) = (A_i, b_i, c_i)] = p_i$, il est immédiat que $E[\varphi(A, b, c)]$ est déterminé par résolution du programme linéaire.

$$A_i x_i \leq b_i \quad i \in I \text{ fini}$$

$$x_i \geq 0$$

$$\text{maximer } \sum_i p_i < c_i, x_i >$$

Un tel programme a l'inconvénient d'être souvent de taille prohibitive, il est alors préférable de procéder de la manière suivante.

DEUXIEME METHODE

Il est toujours possible de définir l'ensemble I des modalités de (A, b, c) comme produit de trois ensemble R, S, T avec $\Pr(A = \bar{A}_r, b = b_s, c = c_t) = p_{rst}$ et c'est souvent sous la forme d'un ensemble factoriel de combinaisons pour (A, b, c) que se poseront les problèmes pratiques correspondant à une hypothèse forte opposée à une hypothèse faible et à une ou plusieurs hypothèses moyennes.

Soit K_r l'ensemble des groupements de m indices de $J = [1, \dots, m+n]$ tel que si $J_0 \in K_r$, $(\bar{A}_r)_{J_0}$ soit non singulière.

Soit S_{J_0} le sous-ensemble de S tel que $(\bar{A}_r)_{J_0}^{-1} b_s \geq 0$ et T_{J_0} le sous-ensemble de T tel que $(\bar{c}_t)_{J_0} (\bar{A}_r)_{J_0}^{-1} \bar{A}_r - \bar{c}_t \geq 0$.

On aura alors :

$$E(\varphi(A, b, c)) = \sum_{r \in R} \sum_{J_0 \in K_r} \sum_{(s, t) \in S_{J_0} \cdot T_{J_0}} (\bar{c}_t)_{J_0} (\bar{A}_r)_{J_0}^{-1} b_s p_{rst}$$

Par le même procédé, on peut trouver les moments de la loi de distribution de $\varphi(A, b, c)$ ou, plus généralement, sa fonction caractéristique.

Si A, b, c sont indépendants en probabilité, les calculs se simplifient.

Bien entendu, il ne peut être question de former K_r à priori et de faire parcourir cet ensemble à J_0 . Il conviendra, au contraire, ayant choisi r, s, t tels que $(\bar{A}_r, b_s, c_t)$ soit voisin de (Ea, Eb, Ec) , de chercher J_0 tel que la base J_0 soit optimale pour le problème

$$\sup [< c_t, x >, A_r x \leq b_s, x \geq 0]$$

puis on formera S_{J_0} et T_{J_0} et on calculera alors

$$U_{r, J_0} = \sum_{(s, t) \in S_{J_0} \cdot T_{J_0}} (\bar{c}_t)_{J_0} (\bar{A}_r)_{J_0}^{-1} b_s p_{rst}; \text{ puis, ayant éliminé dans } S, T \text{ les couples}$$

(s, t) tels que $(s, t) \in S_{J_0} \cdot T_{J_0}$, on passera de J_0 à J'_0 ne différant de J_0 que par un seul indice et tel que $S_{J'} \cdot T_{J'} \neq \phi$. On formera alors la nouvelle quantité U_{r, J'_0} ; le nombre d'éléments restant dans S, T est décroissant; comme il est fini, on formera en un nombre fini d'étapes :

$$U_r = \sum_{J_0 \in K_r} U_{r, J_0}$$

Puis, de la même manière, on formera tous les U_r et, finalement,

$$E [\varphi(A, b, c)] = \sum_{r \in R} U_r$$

Cette façon de procéder permettra, en outre, de vérifier l'hypothèse que le support de la distribution de probabilité est contenu dans Π sous-ensemble de E_d où $\varphi(A, b, c)$ est définie et bornée.

REMARQUE

Cette procédure est bien entendu valable si seule une ou deux composantes A, b, c de (A, b, c) est aléatoire.

Si on a pris soin de calculer en même temps que $m_1 = E[\varphi(A, b, c)]$ le moment 2 : $m_2 = E[\varphi(A, b, c)]^2$, on pourra former $\sigma_2 = m_2 - m_1^2$ et, d'après la relation de Bienaymé Tchebicheff, on aura :

$$\Pr (m_2 - t\sigma \leq \varphi(A, b, c) \leq m_1 + t\sigma) > 1 - \frac{1}{t^2}$$

ce qui apportera certaines précisions quant au support de la distribution de $\varphi(A, b, c)$.

★

★

★

CHAPITRE III

Le chapitre II a été consacré principalement à l'étude de l'élément aléatoire $\varphi(A, b, c)$ et à ce qui pouvait être dit de sa fonction de distribution. C'était donc le point de vue de la prévision qui a été plus particulièrement développé. Le présent chapitre sera au contraire consacré au point de vue de l'action. Deux points de vue seront développés. En premier lieu, on examinera ce qui peut être dit quant au domaine des solutions uniformément admissibles d'un programme linéaire stochastique et, plus généralement, quant au domaine des solutions admissibles avec une probabilité donnée. En deuxième lieu, on exposera deux types de problèmes de décision.

3,1 - Comme précédemment, nous supposons que le support Δ de la distribution de probabilité L de $x = (A, b, c)$ est borné et contenu dans Π domaine d'existence de $\varphi(A, b, c)$ bornée.

On appellera domaine des solutions uniformément admissibles le sous-ensemble de E_n éventuellement vide, tel que :

$$\bar{D} = [x \mid Ax \leq b, x \geq 0, \forall (A, b, c) \in \Delta]$$

Soit $D(A, b, c)$ le polyèdre tel que :

$$D(A, b, c) = [x \mid Ax \leq b, x \geq 0]$$

$D(A, b, c)$ existe et n'est pas vide pour tout $(A, b, c) \in \Delta$.

Le domaine des solutions uniformément admissibles sera :

$$\bar{D} = \bigcap_{(A, b, c) \in \Delta} D(A, b, c)$$

$\bar{D}$ est évidemment convexe, peut-être vide, mais n'est pas en général un polyèdre si la loi L est de type continu. Par contre, si L est une loi discrète finie, $\bar{D}$ est un polyèdre. Si I est le sous-ensemble fini des indices des modalités de (A, b, c) , $\bar{D}$ sera défini par

$$\bar{D} = \bigcap_{i \in I} D_i$$

où encore : $\bar{D} = [x \mid A_i x_i \leq b_i, x \geq 0 \mid \forall i \in I]$

et, avec les mêmes notations qu'au chapitre 2, on aura :

$$D^- \subset \bar{D} \text{ avec } D^- = [x \mid A^+ x \leq b^-, x \geq 0]$$

Le domaine $\bar{D}$ est particulièrement intéressant à connaître au cas où l'on envisage une stratégie prudente.

Si c n'est pas aléatoire, il conviendra de maximiser $\langle c, x \rangle$ dans $\bar{D}$. Si c est aléatoire, on pourra ou bien étudier la fonction de répartition de $\varphi(c)$ dans $\bar{D}$, ou maximiser $\langle E(c), x \rangle$ dans $\bar{D}$.

3,2 - On peut aussi prendre une attitude moins prudente et chercher à déterminer le domaine $\bar{D}_p$ des solutions admissibles avec une probabilité supérieure ou égale à une probabilité donnée p .

Soit G_p l'ensemble des parties Δ_p de Δ telles que :

$$\Pr [(A, b, c) \in \Delta_p] = p$$

On aura alors :

$$\bar{D}_p = \bigcup_{\Delta_p \in G_p} \left(\bigcap_{(A, b, c) \in \Delta_p} D(A, b, c) \right)$$

En général, $\bar{D}_p$ ne sera pas convexe, (sauf $\bar{D}_1 = \bar{D}$) même si la loi est discrète finie.

On aura $\bar{D} \subset \bar{D}_p \quad \forall p \in [0, 1]$ et en particulier $\bar{D}_0 = \bigcup_{(A,b,c) \in \Delta} D(A,b,c)$,

$\bar{D}_0$ est le domaine des solutions possibles. On peut se proposer de maximiser $\langle E(c), x \rangle$ dans $\bar{D}_p$ (ou $\langle c, x \rangle$ si c n'est pas aléatoire). Ce problème est soluble dans le cas où L est une loi discrète finie à l'aide d'un programme mixte faisant intervenir, outre la variable x continue, des variables booléennes.

En effet, soient :

- I l'ensemble fini des modalités de (A,b,c) avec $p_i = \Pr [(A,b,c) = (A_i, b_i, c_i)]$
- J_p le sous-ensemble des parties I_p de I telles que $\sum_{i \in I_p} p_i \geq p$

Autrement dit :

$$J_p = \{ I_p \mid I_p \subset I, \sum_{i \in I_p} p_i \geq p \}$$

On peut alors écrire $\bar{D}_p$ sous la forme :

$$\bar{D}_p = \bigcup_{I_p \in J_p} D_{I_p} = \bigcup_{I_p \in J_p} \left[\bigcap_{i \in I_p} D_i \right]$$

$\bar{D}_{I_p}$ s'écrit encore :

$$\bar{D}_{I_p} = \{ x \mid A_i x \leq b_i, x \geq 0, \forall i \in I_p \}$$

ou enfin, en introduisant les matrices :

$$A_{I_p} = [A_i \mid i \in I_p] \text{ de format } (|I_p| \cdot m) \cdot n$$

$$b_{I_p} = [b_i \mid i \in I_p] \text{ de format } (|I_p| \cdot m) \cdot 1$$

$$\bar{D}_{I_p} = \{ x \mid A_{I_p} x \leq b_{I_p}, x \geq 0 \}$$

Soit M le cube fermé de R^n contenant $\bar{D}_p$ de longueur d'arête ℓ :

$$M = \{ x \mid 0 \leq x_i \leq \ell \}$$

On peut considérer $\bar{D}_{I_p}$ pour un certain $I_p \in J_p$ comme l'intersection des demi-espaces fermés H_j définis par $f_j(x) \leq 0$ où $f_j(x)$ est la forme affine:

$$f_j(x) = A_{I_p,j} x - b_j = \sum_{i=1}^n A_{I_p,j}^i x_i - b_j$$

obtenue en notant $A_{I_p,j}$ la $i^{\text{ème}}$ ligne de A_{I_p} et b_j la $j^{\text{ème}}$ composante de b_{I_p} .

Dans le cube M , $f_j(x)$ a un maximum m_j tel que :

$$m_j = \ell \sum_{i=1}^n \sup (0, A_{I_p,j}^i) - b_j$$

Soit alors m_{I_p} la matrice de format $(|I_p| \cdot m) \times 1$ formée par les m_j .

Le problème à résoudre sera alors :

maximer $\langle E(c), x \rangle$
 compte tenu de :

$$\begin{array}{ll}
 (1) & A_{Ip} x - b_{Ip} - m_{Ip} \xi_{Ip} \leq 0 \\
 (2) & \xi_{Ip} \text{ variable booléenne (0 ou 1)} \\
 (3) & \sum_{Ip \in J_p} \xi_{Ip} = |J_p| - 1 \\
 (4) & x \geq 0
 \end{array} \quad \forall Ip \in J_p$$

En effet, d'après (3), toutes les variables booléennes ξ_{Ip} seront égales à 1, sauf une, donc tous les groupes de relations (1) seront identiquement vérifiés sauf un, d'après la définition de m_{Ip} . L'élément $\bar{x}$ solution sera point extrême du polyèdre $\bar{D}_{Ip}$ tel que $\xi_{Ip} = 0$ et la probabilité que l'élément x soit admissible sera exactement $p' = \sum_{i \in Ip} p_i \geq p$.

REMARQUE

L'introduction des variables booléennes ξ a permis de définir indirectement la fermeture convexe de l'ensemble des profils $\bar{D}_{Ip}$ de tous les D_{Ip} .

La résolution d'un tel problème mixte peut être effectivement réalisée par une méthode (non publiée) que nous avons mise au point.

3.3. $\bar{D}_p$ se présente comme réunion finie (dans le cas où la loi L est discrète finie) de polyèdres D_{Ip} , $Ip \in J_p$.

Si Ip est donné dans J_p , la probabilité que $x \in D_{Ip}$ soit réalisable est exactement $p'_{Ip} = \sum_{i \in Ip} p_i \geq p$

Soit H_k le demi-espace fermé de R^n tel que $\langle Ec, x \rangle \geq k$.

On peut se poser le problème de chercher le sous-ensemble $\Delta_{Ip,k} = \bar{D}_{Ip} \cap H_k$ tel que, si $x \in \Delta_{Ip,k}$, la probabilité p'_{Ip} que x soit réalisable soit maxima.

L'intérêt d'un tel domaine $\Delta_{Ip,k}$ est évident car tous les éléments de ce $\Delta_{Ip,k}$ sont réalisables avec une probabilité maximum parmi tous les éléments de $\bar{D}_p$ qui conduisent à $\langle Ec, x \rangle \geq k$.

METHODE DE RESOLUTION

Supposons connu le sous-ensemble J_{pk} de J_p tel que

$$J_{pk} = [Ip \mid Ip \in J_p, \bar{D}_{Ip} \cap H_k \neq \emptyset]$$

le sous-ensemble $\bar{I}_p$ cherché sera déterminé par :

$$\sum_{i \in \bar{I}_p} p_i = \sup_{Ip \in J_{pk}} \left(\sum_{i \in Ip} p_i \right)$$

et on aura :

$$\Delta_{Ip,k} = D_{\bar{I}_p} \cap H_k$$

Pour déterminer $J_{p,k}$ il suffira de chercher I_p tel que $\Delta_{I_p,k} = D_{I_p} \cap H_k \neq \emptyset$ ce qui conduit à résoudre une suite finie de programmes linéaires (en effet, $\Delta_{I_p,k}$ est un polyèdre de R^n ; pour voir s'il est vide ou non, il suffit de chercher à maximiser dans $\Delta_{I_p,k}$ n'importe quelle fonction économique).

REMARQUE

Sur le plan pratique, de telles procédures ne seront réalisables que si $|J_p|$ n'est pas trop élevé.

3,4 - Mentionnons, pour terminer ce chapitre, deux types de problèmes se présentant comme des problèmes de décision.

3,41 - Le premier type de problème étudié par W. Krelle se pose de la façon suivante :

Supposons avoir choisi à priori un vecteur $x \geq 0$ caractérisant une certaine stratégie (par exemple Une stratégie de production pour une entreprise de production) avant la réalisation de l'évènement (A_i, b_i, c_i) qui est supposé se produire avec une probabilité p_i suivant une loi L discrète et finie à support borné dans Π .

Si (A_i, b_i, c_i) se produit effectivement, il peut être avantageux ou obligatoire (si $A_i x \leq b_i$ n'est pas vérifié) de choisir à posteriori une stratégie de « repli » x_i telle que $x_i - x \leq 0$. Mais, cette modification de stratégie présente un coût d'écart $+ < e, x - x_i >$.

Le problème posé est : comment choisir la stratégie x à priori pour que l'espérance mathématique de la fonction économique totale (y compris les coûts d'écart) soit maximum.

Remarquons que, si l'évènement i se produit, la fonction économique prend la valeur :

$$y_i = < c_i, x_i > - < e, x - x_i >$$

Par suite, l'espérance de la fonction économique est :

$$\varphi = \sum_{i \in I} p_i y_i = \sum_i p_i < c_i + e, x_i > - < e, x >$$

Il conviendra donc de choisir x tel que φ soit maximum, compte tenu des contraintes :

$$A_i x_i \leq b_i \quad i \in I$$

$$x_i - x \leq 0$$

$$x_i \geq 0$$

$$x \geq 0$$

C'est un programme linéaire classique qui donnera comme solutions à la fois la stratégie « à priori » x et les stratégies « à posteriori » x_i .

3,42 - G.B. Dantzig, A. Fulkerson et Madansky ont étudié des types de problèmes de décision à plusieurs étages. Pour le principe, il suffit d'envisager un problème à 2 étages. Ce problème est le suivant.

En premier lieu, on choisit x tel que :

$$Ax \leq b_1 \quad x \geq 0$$

A, b_1 étant deux matrices non aléatoires, puis x étant choisi, le hasard détermine une matrice b_2 ; alors, on détermine y tel que :

$f(x, y) = \langle u, x \rangle + \langle v, y \rangle$ soit maximum compte tenu de $Bx + Cy = b_2, y \geq 0$

Comment choisir x «à priori» tel que $E(f(x, y))$ soit maximum?

Supposons que les modalités de b_2 soient discrètes et finies, et soit p_i la probabilité de $b_2 = b_{2i}$ suivant une loi à support dans Π .

Soit $\bar{y}_i$ tel que

$$\langle u, x \rangle + \langle v, \bar{y}_i \rangle = \sup [\langle c, x \rangle + \langle v, y \rangle \mid Bx + Cy_i = b_{2i}, y_i \geq 0]$$

avec x valeur fixée à priori tel que : $Ax \leq b_1, x \geq 0$.

On aura alors :

$$E[f(x, y)] = \langle u, x \rangle + \sum_{i \in I} p_i \langle v, \bar{y}_i \rangle$$

Le problème se ramène donc au programme linéaire suivant :

$$\text{maximer } \langle u, x \rangle + \sum_{i \in I} p_i \langle v, y_i \rangle$$

compte tenu de :

$$Ax \leq b_1$$

$$Bx + Cy_i = b_{2i} \quad i \in I$$

$$x_i \geq 0, \quad x \geq 0$$

Ce problème peut encore être résolu de la manière suivante :

Les relations $Ax \leq b_1, x \geq 0$ déterminent un domaine D_1 dans E_n .

Soit x fixé dans D_1 , alors y doit maximiser $\langle v, y \rangle$ sous les contraintes

$$Cy = b_2 - Bx$$

$$y \geq 0$$

b_2 étant une aléatoire suivant une loi L .

Supposons que pour tout $x \in D_1$ et pour tout b_2 suivant la loi L $(C, b_2 - Bx, v)$ appartienne au domaine Π où le maximum de $\langle v, y \rangle$ sous les contraintes précédentes est défini et borné.

Le problème dual du problème précédent est

$$\text{minimer : } \langle z, b_2 - Bx \rangle$$

compte tenu de $zC \geq v$

C'est un problème stochastique particulier ou seule la forme à maximiser a ses composantes aléatoires.

Soit K l'ensemble des indices des structures S_k correspondant au problème,

$\sup [\langle l, z \rangle \mid zC \geq v]$ quand seule l est susceptible de varier. Tant que $l \in S_k$ le point extrême optimal Z_k du problème précédent ne dépend pas de l ,

si $l = b_2 - Bx \in S_k$ on a :

$$\sup [\langle v, y \rangle] = \langle b_2 - Bx, Z_k \rangle$$

Si b_2 suit une loi L de distribution (continue par exemple) $f(b_2)$, la loi de l sera $f(1+Bx)$ et on aura alors :

$$\begin{aligned} E(\langle v, y \rangle) &= \sum_{k \in K} \int_{S_k} \langle l, Z_k \rangle f(1+Bx) dl \\ &= \sum_{k \in K} \langle \lambda_k(x), Z_k \rangle = H(x) \end{aligned}$$

Si, maintenant, x n'est plus fixé dans D_1 , le problème revient à trouver x tel que :

$$Ax \leq b_1$$

$$x \geq 0$$

et maximant $\langle u, x \rangle + H(x)$

Ce problème n'est pas linéaire car, en général, H n'est pas une forme linéaire mais il est de taille sensiblement plus réduite que le premier. G.B. Dantzig a, par ailleurs, démontré que $H(x)$ était une fonction concave. Un tel problème pourra donc se résoudre par exemple à l'aide de la méthode du gradient itéré.

★

★

★

BIBLIOGRAPHIE

La bibliographie sur la programmation linéaire par méthodes simpliciales ne sera pas donnée parce que classique et abondante (voir par exemple la bibliographie indiquée dans M. COURTILLOT, Thèse d'Ingénieur Docteur, Paris 1959).

OUVRAGES GENERAUX

- N. BOURBAKI Algèbre et Topologie, Hermann Paris
- C. BERGE Espaces topologiques et fonctions multivoques, Dunod Paris
- CRAMER Mathematical methods of statistics, Princeton University Press
- ARROW-HURWICZ-UZAWA Studies in linear and nonlinear programming, Standford University Press
- Linear inequalities and related systems, Princeton University Press.
- PONTRYHAGIN Foundations of combinatorial Topology.

REFERENCES SPECIALISEES

1) Méthodes non classiques de résolution des programmes linéaires et de résolution de programmes linéaires

- J.F. BENDERS Partitioning in Mathematical Programming. Thèse de Doctorat. Université d'Utrecht.
- RAGNAR FRISCH Principle of linear programming - Memorandum fra Universitetets Sosialokonomiske Institutt Oslo
- J.B. ROSEN Gradient projection method for non linear programming Document interne Shell Developpement Company
- G. PARISOT Les programmes logarithmiques, leur application au calcul des programmes linéaires.
C.R.A.S. Tome 248 p. 1470-1472 (1959) et Revue Française de Recherche Opérationnelle N° 20 (1961).
- M. PEUCHOT Méthode des potentiels maxima,
Présentée au Congrès International de Recherche Opérationnelle, Aix-en-Provence 1960.

2) Programmation stochastique

- R. FORTET Conférence AFCAL TI-SOFRO, Groupe MPE 5/6/61.
- E.M.L. BEALE On minimising a convex function subject to linear inequalities.
Journal of the Royal Statistical Society, série B, vol. 17 N° 2 1955.
- G.B. DANTZIG Linear Programming under uncertainty
RM 1374-1 Rand Corporation 1955
- A. MADANSKY Inequalities for statistic linear programming problems. Management Science vol. 6 N° 2 (1960)
- A. MADANSKY Some results and problems in stochastic linear programming. p. 1596
Rand Corporation 1953
- W. KRELLE Linear programming under uncertainty
Papier de l'European Meeting of the Econometric Society. Amsterdam 1959.
- A. CHARNES and COOPER Chance constrained programming
Management Sciences vol. 6 N° 1 1959
- S.E. ELMAGHRABY An approach to linear programming under uncertainty
ORSA Vol. 7 N° 2 1959
- L. HURWICZ Stochastic Models of econometric fluctuations
Econometrica Vol. 12 N° 2 1944.

TABLE DES MATIERES

INTRODUCTION	Pages
PREMIERE PARTIE	
- CHAPITRE I : DEFINITIONS ET PRINCIPAUX THEOREMES	1
1,1 - Variétés linéaires - Cônes - Ensembles convexes	1
1,2 - Introduction d'une métrique	4
1,3 - Polyèdres	6
1,4 - Systèmes d'inégalités linéaires	8
- CHAPITRE II : LA METHODE DITE DU SIMPLEXE	13
2,1 - Généralités	13
2,2 - La méthode simplexe :	16
. Méthode directe	16
. Méthode duale	18
. Interprétation en termes de géométrie affine	20
- CHAPITRE III : LA METHODE DU DOUBLE GRADIENT DE M. RAGNAR FRISCH.....	30
3,1 - Généralités	30
3,2 - Méthode de résolution	30
3,3 - Essai de résolution d'un programme linéaire par intégration d'équation différentielles dérivant de la méthode de M. R. Frisch	39
3,4 - Méthodes se rattachant à la méthode du double gradient	41
- CHAPITRE IV : LA METHODE SEQUENTIELLE INTERIEURE	43
4,1 - Généralités et principe de la méthode	43
4,2 - Construction de la suite	44
4,3 - Procédé pratique de calcul	44
4,4 - Variante de la méthode séquentielle intérieure	52
- CHAPITRE V : LA METHODE DU GRADIENT ITERE	56
5,1 - Généralités	56
5,2 - Méthode de calcul	60
- CHAPITRE VI : GENERALISATION DE LA METHODE DU GRADIENT ITERE A UNE CLASSE DE PROGRAMME NON LINEAIRE	64
6,1 - Généralités	64
6,2 - Extension de la méthode du gradient itéré	65
6,3 - La méthode du gradient projeté	66

DEUXIEME PARTIE : PROGRAMMATION LINEAIRE STOCHASTIQUE

Pages

- CHAPITRE I	71
1,1 - Points de vue possibles	71
1,2 - Relations entre problèmes stochastiques et problèmes paramétrés	72
- CHAPITRE II	78
2,1 - Généralités	78
2,2 - Majorations	78
2,3 - Calcul direct - Détermination des fonctions caractéristiques	83
- CHAPITRE III	87
3,1 - Détermination du domaine des solutions uniformément admissibles	87
3,2 - Détermination du domaine D_p	87
3,4 - Deux problèmes de décision	90
- Bibliographie	
- Errata	

DEUXIEME THESE

Propositions données par la Faculté

« POINTS RATIONNELS SUR LES CUBIQUES »

Vu et approuvé :

Paris, le 19 Avril 1963

Le Doyen de la Faculté des Sciences,

Marc ZAMANSKY

Vu et Permis d'imprimer :

Le Recteur de l'Académie de Paris,

Jean ROCHE

